MONTHLY NOTICES

OF THE

ROYAL ASTRONOMICAL SOCIETY

Volume 121 No. 2 1960

Published and Sold by the

ROYAL ASTRONOMICAL SOCIETY
BURLINGTON HOUSE
LONDON, W.1

Price £1 4s. 0d.; *in U.S.A.* $3.50

Subscription for volume: £6 ; *in U.S.A.* $18

The Geophysical Journal

OF THE

ROYAL ASTRONOMICAL SOCIETY

Editors

A. H. Cook
M.A., Ph.D., F.R.A.S., F.G.S.
National Physical Laboratory
Teddington

T. F. Gaskell
M.A., Ph.D., F.R.A.S.
British Petroleum Company
London

Price £1 per number; in U.S.A. $3. Annual Subscription £3; in U.S.A. $9

Volume 3 No. 3 September 1960

CONTENTS INCLUDE :

Orders should be addressed to :

The Assistant Secretary

Royal Astronomical Society, Burlington House, London, W.1

NOTICE TO AUTHORS

Presentation of Papers at Meeting

At some meetings of the Society the background and conclusions of selected papers are presented and then discussed. In order to assist the Secretaries in the selection of papers for such meetings, authors are asked to let the Society know, when submitting papers, whether they would be willing to give an account of their paper, if requested.

The attention of authors resident abroad is drawn to the fact that the Society welcomes information about their work. The Secretaries would be happy to consider having such work described at a meeting, in accordance with the author's wishes, either by a Secretary or other Fellow.

Publication of Papers

1. *General.*—It is the aim of the Society to be of the greatest possible service in disseminating astronomical results and ideas to the scientific community with the utmost possible speed. Contributors are accordingly urged to give the most careful consideration to the presentation of their work, for attention to detail will assuredly result in a substantial saving of time.

It is the practice of the Society to seek a referee's opinion on nearly every paper submitted for publication in *Monthly Notices*; experience has shown that frequently the comments of referees have enabled authors to improve the presentation of their work and so increase its scientific value.

2. *Communication.*—Papers must be communicated to the Society by a Fellow. They should be accompanied by a summary at the *beginning* of the paper conveying briefly the content of the paper, and drawing attention to important new information and to the main conclusions. The summary should be intelligible in itself, without reference to the paper, to a reader with some knowledge of the subject; it should not normally exceed 200 words in length. **Authors are requested to submit MSS. in duplicate.** These should be typed using double spacing and leaving a margin of not less than one inch on the left-hand side. Corrections to the MSS. should be made in the text and not in the margin. By Council decision, MSS. of accepted papers are retained by the Society for one year after publication; unless their return is then requested by the author they are destroyed.

3. *Presentation.*—Authors are allowed considerable latitude, but they are requested to follow the general style and arrangement of *Monthly Notices*. References to literature should be given either in the traditional form of a numbered list at the end of the paper, or as prescribed in *Notes on the Preparation of Papers to be Communicated to the Royal Society*.

4. *Notation.*—For technical astronomical terms, authors should conform closely to the recommendations of Commission 3 of the International Astronomical Union (*Trans. I.A.U.*; Vol. VI, p. 345, 1938). Council has decided to adopt the I.A.U. 3-letter abbreviations for constellations where contraction is desirable (Vol. IV, p. 221, 1932). In general matters, authors should follow the recommendations in *Symbols, Signs and Abbreviations* (London: Royal Society, 1951) except where these conflict with I.A.U. practice.

5. *Diagrams.*—**These should be designed to appear upright on the page, drawn about twice the size required in print and prepared for direct photographic reproduction except for the lettering, which should be inserted in pencil.** Legends should be given in the manuscript indicating where in the text the figure should appear. Blocks are retained by the Society for 10 years; unless the author requires them before the end of this period they are then destroyed. **Rough copies or prints of the diagrams should accompany each manuscript.**

6. *Tables.*—These should be arranged so that they can be printed upright on the page.

7. *Proofs.*—Authors are liable for costs of alteration exceeding 5 per cent of composition. It is therefore in their own and the Society's interests to seek the maximum conciseness and simplification of symbols and equations consistent with clarity.

CONTENTS

Printed in London by Taylor & Francis, Ltd., 1960 *August*

MONTHLY NOTICES
OF THE
ROYAL ASTRONOMICAL SOCIETY
Vol. 121 No. 2

THE NEW I.A.U. SYSTEM OF GALACTIC COORDINATES (1958 REVISION)

(PAPER 1)

A. Blaauw, C. S. Gum, J. L. Pawsey and G. Westerhout*

(Members of the I.A.U. Sub-Commission 33b)

(Received 1960 March 21)

Summary

The definition of a new system of galactic coordinates was recently announced by Sub-Commission 33b on behalf of the I.A.U. The present paper is the first of a series of five, which together form the final report of the Sub-Commission. It summarizes the observational evidence, given in detail in the other papers, relevant to setting up the new system, it discusses various considerations which were taken into account, and it concludes with a definition of the new system.

1. *Introduction*

The definition of a new system of galactic coordinates was recently announced by the I.A.U. Sub-Commission 33b on behalf of the International Astronomical Union.

The present series of papers is the detailed and final report of the Sub-Commission. The various investigations, decisions, and recommendations of both the Sub-Commission and the 1958 Moscow General Assembly are presented in general terms in this paper (Paper I), which is a coordinating paper. It is followed by four more detailed papers, prepared on behalf of the Sub-Commission, which discuss specific aspects of the investigation. These are:

Paper II. *A 21-cm determination of the principal plane of the Galaxy*, by C. S. Gum, F. J. Kerr and G. Westerhout.

Paper III. *Radio data relevant to the choice of a galactic coordinate system*, by C. S. Gum and J. L. Pawsey.

Paper IV. *Optical determinations of the galactic pole*, by A. Blaauw.

Paper V. *The position of the galactic centre*, by J. H. Oort and G. W. Rougoor.

Of the detailed papers, Paper II is an original analysis of combined Leiden and Sydney observations. The others present compilations of published and unpublished material and its bearing on the selection of a coordinate system. In addition, the papers bring out a number of physical relationships of general interest to the problem of galactic structure, such as the close relationship in the galactic disk between the distributions of the sources of the 21-cm radiation and of the radio continuum.

2. *Historical*

William Herschel in 1785 first realized the importance for investigations of the Milky Way System of a special coordinate system having its equator in the galactic plane, and from his time until 1932 various systems having poles

*Dr. Gum was unfortunately killed in a skiing accident in Switzerland on 1960 April 28, after the completion of this group of papers.

CONTENTS

Printed in London by Taylor & Francis, Ltd., 1959 August

MONTHLY NOTICES
OF THE
ROYAL ASTRONOMICAL SOCIETY
Vol. 121 No. 2

THE NEW I.A.U. SYSTEM OF GALACTIC COORDINATES (1958 REVISION)

(PAPER I)

A. Blaauw, C. S. Gum, J. L. Pawsey and G. Westerhout*

(Members of the I.A.U. Sub-Commission 33b)

(Received 1960 March 21)

Summary

The definition of a new system of galactic coordinates was recently announced by Sub-Commission 33b on behalf of the I.A.U. The present paper is the first of a series of five, which together form the final report of the Sub-Commission. It summarizes the observational evidence, given in detail in the other papers, relevant to setting up the new system, it discusses various considerations which were taken into account, and it concludes with a definition of the new system.

1. *Introduction*

The definition of a new system of galactic coordinates was recently announced by the I.A.U. Sub-Commission 33b on behalf of the International Astronomical Union.

The present series of papers is the detailed and final report of the Sub-Commission. The various investigations, decisions, and recommendations of both the Sub-Commission and the 1958 Moscow General Assembly are presented in general terms in this paper (Paper I), which is a coordinating paper. It is followed by four more detailed papers, prepared on behalf of the Sub-Commission, which discuss specific aspects of the investigation. These are:

Paper II. *A 21-cm determination of the principal plane of the Galaxy*, by C. S. Gum, F. J. Kerr and G. Westerhout.

Paper III. *Radio data relevant to the choice of a galactic coordinate system*, by C. S. Gum and J. L. Pawsey.

Paper IV. *Optical determinations of the galactic pole*, by A. Blaauw.

Paper V. *The position of the galactic centre*, by J. H. Oort and G. W. Rougoor.

Of the detailed papers, Paper II is an original analysis of combined Leiden and Sydney observations. The others present compilations of published and unpublished material and its bearing on the selection of a coordinate system. In addition, the papers bring out a number of physical relationships of general interest to the problem of galactic structure, such as the close relationship in the galactic disk between the distributions of the sources of the 21-cm radiation and of the radio continuum.

2. *Historical*

William Herschel in 1785 first realized the importance for investigations of the Milky Way System of a special coordinate system having its equator in the galactic plane, and from his time until 1932 various systems having poles

*Dr. Gum was unfortunately killed in a skiing accident in Switzerland on 1960 April 28, after the completion of this group of papers.

differing from each other by one or two degrees were used. The literature up to 1932 is summarized in the Introduction to the well-known *Lund Observatory Conversion Tables* prepared by Ohlsson (1932). For these Tables the pole at right ascension $12^h 40^m$, declination $+28°$ (1900·0), used by Pickering for his extensive researches at Harvard Observatory, was adopted by Ohlsson, and the zero point of longitude was taken at the intersection of the galactic plane and the celestial equator for the equinox 1900.0. The system of galactic coordinates thus defined became the standard one and we shall hereafter refer to it as the "old" galactic coordinate system. The recent *Lund Observatory Tables for the Conversion of Galactic into Equatorial Coordinates for the Epoch* 1958.0 (Ohlsson, Reiz and Torgård 1956) are also based on the old system.

The most extensive optical investigation since 1932 is that of van Tulder (1942). The mean position of the pole finally adopted by van Tulder deviated from the Ohlsson pole by about one degree. However, the need for a new pole was probably not very strongly felt at this time owing to the fact that the scatter of the poles derived from various groups of objects was of a similar order to, or larger than, the deviation itself. Perspectives began to change with the advent of radio surveys carried out at various frequencies during the last ten years or so. Almost without exception these surveys showed that, in the vicinity of the galactic centre, the maximum of the radio emission fell south of the old equator by something over a degree. The deviation is in the same sense as that found by van Tulder; in fact, the optical data of van Tulder, and also those examined more recently by Ashbrook and Duncombe (1952) and by Kirillova (1955), are in much closer agreement with the radio data than with the old pole.

On the basis of this type of evidence it has been thought for some time that a revision of the galactic coordinate system should be considered, and the International Astronomical Union at its General Assembly in Dublin in 1955 appointed Sub-Commission 33b "to investigate the desirability of a revision of the position of the galactic pole and of the zero point of galactic longitude".

The Sub-Commission's investigations had, by the time of the 1958 Moscow General Assembly, reached the stage where it could recommend that the time was opportune for the adoption of a new system of galactic coordinates. The Sub-Commission also formulated the principles which it considered should be used in selecting positions for the new pole and zero of longitude. There remained, however, some further study before the actual positions could be selected.

On the basis of these recommendations, supplemented by detailed reports which may be regarded as preliminary drafts of the papers in this present series, the General Assembly passed a resolution endorsing the above recommendations and authorizing the Commission to define and announce the new system as soon as its studies were completed. In March 1959 the Commission completed its investigations, decided on the definition of the new system and communicated its decision to the General Secretary of the I.A.U. and to a number of astronomical journals throughout the world (Blaauw *et al.* 1959).

We shall present the considerations leading to the adoption of the new system by giving in Sections 3 and 4 the substance of the report of the Commission to the Moscow General Assembly, and in Section 5 the resolution passed by the General Assembly. Section 6 describes the considerations leading to the subsequent choice of the actual pole and zero of longitude. We conclude with extracts from

the text of the announcement defining the new system, which was communicated by the Commission to the General Secretary of the I.A.U.

3. *Evidence on the locations of the pole and centre of the Galaxy*

In examining the basis for a possible revision of the galactic coordinate system it was necessary first to determine the position, with respect to the Sun, of the galactic plane and of the centre of the Galaxy, and to estimate the uncertainties in these determinations. Any decision on the adoption of a new system depended on considerations such as the extent to which the old pole was misplaced from the true one, in relation to the current uncertainty or to expected improvements in the foreseeable future. Both optical and radio observations were examined but it turned out that radio ones were the more significant.

3.1. *The position of the pole.*—Evidence on the precise location of the galactic pole is based on the study of objects strongly concentrated towards the galactic plane. The relevant observations consist of (*a*) radio observations of the 21-cm line emitted by neutral interstellar hydrogen (H I regions); (*b*) radio observations of the continuum, mainly in the wavelength range 20 cm to 3·5 m; and (*c*) optical observations.

Of these, the effective optical observations are restricted to a relatively small region around the Sun, about 3000 pc in radius, i.e. to an area of 5 or 10 per cent of that accessible to radio observations; the objects concerned represent less than $\frac{1}{10}$ per cent of the total mass of the Galaxy. Observations of the 21-cm line pertain to the neutral hydrogen, which represents about 2 per cent of the total mass of the Galaxy, and which can be located over the greater part of the galactic plane out to a radius of about 15 kpc from the galactic centre. The radio continuum observations give only integrated effects along the line-of-sight and are most effectively used in checking a model based on other data.

The outstanding feature of the whole of the observations is the high degree of flatness of the layer of neutral hydrogen in the region of the Galaxy within about 7 kpc* of the centre (see Paper II). Within this area nearly all the points of maximum density of hydrogen in the direction normal to the galactic plane (smoothed over areas of about 1 kpc^2) lie within 20 pc of the mean plane of the layer (which is described as the "H I Principal Plane"). These deviations are less than 1 part in 700 of the diameter of the region. In the outer regions of the Galaxy systematic distortions of up to 800 pc occur (see Fig. 1, Paper II, p. 133). The high degree of flatness of the disk of interstellar gas in the inner region must be of fundamental physical significance and the H I principal plane appears to be the best reference plane for dynamical studies of the whole Galaxy. It thus seems reasonable to choose for the equatorial plane of the galactic coordinate system the plane through the Sun parallel to the H I principal plane. Available observations locate the corresponding galactic pole about 1°·5 from the old pole, with an uncertainty of about 0°·1.

Optical observations (see Paper IV) of different groups of objects give poles which are scattered over a range of several degrees and an optical determination of the pole depends on the choice of the group, or groups, of objects to be given the most weight. We have thought it proper to give the greatest weight to observations of the more distant objects (e.g. Cepheids, OB-stars, and open

* Throughout this series of papers the value $R_0 = 8{\cdot}2$ kpc has been adopted for the distance to the galactic centre. The large uncertainty in the absolute distance scale is not important in considering the choice of a galactic coordinate system.

clusters, at distances of about 3000 pc). This weighting corresponds to the assumption that, in the vicinity of the Sun, local irregularities and effects of selection are more serious than possible large-scale distortions. It is supported by the greater consistency of the results from the more distant objects. These selected optical data give consistent agreement with the H I pole to better than 1° for different optical groups (see Fig. 3, Paper IV, p. 168). Such a selection of a small percentage of the optical objects is indeed arbitrary, but the real difficulty is that it is impossible to observe optically a sample of Population I objects which is a representative sample of the whole Galaxy.

In the radio continuum results (see Paper III) the important feature for our purpose is that, in the range of longitudes within about $\pm 60°$ of the galactic centre, where the H I layer is flat, the "ridge lines" of the continuum contour maps are very well defined and agree closely with one another and with the position of the H I layer. In directions where the H I layer is distorted, the continuum ridge line is ill-defined and appears to conform with the H I irregularities. The inference is that, on a large scale, the sources of the "disk component" of the continuum are distributed similarly to the H I regions. The agreement in the flat inner region is within the observational uncertainty of about $0°{\cdot}1$. The continuum observations therefore give strong support to a pole based on the H I observations.

3.2. *The height of the Sun above the plane.*—The distance z_0 of the Sun from the H I principal plane does not appear explicitly in the parameters of a galactic coordinate system, since the required system is necessarily based on a galactic plane through the Sun. But, even if the equatorial plane of the coordinate system were exactly parallel to the H I principal plane, objects in the latter plane will not be on the galactic equator if z_0 differs from zero. The distance of an object from the H I principal plane—presumably the plane of symmetry of the inner parts of the Galaxy—can only be determined when both its distance from the Sun and the value of z_0 are known. For physical reasons it is therefore important to know z_0. Fortunately, as found in Paper II, the Sun is in the H I principal plane within the errors of measurement, the derived value being $z_0 = +4\,\text{pc} \pm 12$ (estimated p.e.). It should be borne in mind in considering observational data, however, that an uncertainty of 12 pc in z_0 corresponds to an uncertainty of $0°{\cdot}09$ in the direction towards a point in the H I principal plane at a distance of 8 kpc from the Sun (e.g. the galactic centre).

The difference between the value of z_0 found from 21-cm line data, $+4\,\text{pc} \pm 12$ (Paper II), and the height of the Sun above the mean plane of the optically observed Population I objects, $+22\,\text{pc} \pm 2$* (Paper IV), need not imply that the mean plane of these objects differs from the H I principal plane throughout the whole Galaxy. The optical region contains an inadequate sample of the whole plane. Moreover the sample could well be biased because a large proportion of the region on which the optical determination is based is situated in those parts of the Galaxy where the neutral hydrogen disk is distorted.

3.3. *The longitude of the galactic centre.*—The (old) longitude of the galactic centre derived from various optical and radio studies (Papers III and V) is $l^{\text{I}} = 327°{\cdot}7 \pm 0°{\cdot}1$ (estimated p.e.). A more precisely determined position is that of the radio source Sagittarius A (17S2A) at $l^{\text{I}} = 327°{\cdot}68$, $b^{\text{I}} = -1°{\cdot}45$ (Paper III), which agrees with the previous direction to within $\frac{1}{3}$ of the

* Note that this error is an internal probable error and would underestimate the uncertainty if the data were subject to systematic bias.

observational uncertainty. We shall, on the basis of evidence presented in Paper V, assume that Sagittarius A is located at the galactic centre. However, even if this assumption later proved to be wrong, its adoption will not have appreciably degraded our placing of the zero of longitude.

4. *Arguments for setting up a new galactic coordinate system now*

The old system of galactic coordinates does not serve its purpose as well as it could because its pole is displaced from the pole of the H I principal plane and from that of Population I objects near the Sun by about 1°·5. Further, the longitude zero is in an inconvenient position, about 32° from the direction of the galactic centre. But if a new system is to be adopted now it may not be perfect for future refined studies. The Sub-Commission recommended a new system on the following grounds:

(1) The old pole is sufficiently seriously misplaced to make the discrepancy obvious for many observations (particularly radio ones). Hence observers might well use their own systems if the I.A.U. did not recommend a new one.

(2) The anticipated improvement in the position of the pole, by a factor of 10/1 or better, is substantial, and is unlikely to be much increased for many years.

(3) A galactic coordinate system is sufficiently accurately placed when the scatter of the objects or groups of objects considered is much greater than the uncertainty in the position of the plane due to observational errors, or due to poor sampling of the objects from which the plane is determined. The proposed system appears to satisfy this condition for the great majority of both optical and radio objects.

(4) This is a time in the history of astronomy when a great increase in the number of studies involving the use of galactic coordinates is anticipated. Consequently the introduction of a new system now, rather than in a few years, should avoid a considerable amount of conversion between the old and the new coordinate systems.

5. *The resolution adopted by the* 1958 *General Assembly*

The Sub-Commission, at a joint meeting of Commissions 33 and 40, recommended the adoption in the near future of a new galactic coordinate system. It recommended a latitude and longitude system similar to the old one, but with the zero of longitude in the direction of the galactic centre, since this direction is now known with reasonable accuracy. This arrangement facilitates studies of the degree of symmetry of the Galaxy about the meridian plane through the galactic centre. Indeed, to further this objective, the Sub-Commission tentatively recommended a longitude system like the terrestial one with positive and negative longitudes, but it was decided at the Moscow meeting that the 0–360° system was preferable.

This meeting passed a resolution with the following clauses, which was subsequently endorsed by the General Assembly:

(*a*) That a standard system of galactic coordinates be adopted for which the pole is based primarily on the distribution of neutral hydrogen in the inner parts of the galactic system.

(*b*) That the zero of longitude be chosen near the longitude of the galactic nucleus, that longitude be counted from 0° to 360°, in the same direction as in the current system, and that latitude be counted in the conventional manner from −90° through 0° to +90°.

(*c*) That Commission 33b be authorized to define the exact values of the coordinates of the pole and of the zero of longitude immediately after the final reduction of the relevant observations is finished.

(*d*) That Commission 33b be charged with the communication of these values to the members of the I.A.U. and to all interested institutions and individuals.

(*e*) That Commission 33b be charged with the supervision of the publication of tables, accurate to 0°·01, necessary for the conversion from galactic into equatorial coordinates and vice versa, and from coordinates based on the Lund pole into the newly defined system and vice versa, and of two conversion charts.

(*f*) That the new galactic coordinates be designated by the same symbols as the old galactic coordinates.

(*g*) That the revised system be referred to as galactic coordinates (1958 revision).

6. *Choice and specification of the final coordinate system*

Following the adoption of this resolution by the I.A.U., the Sub-Commission completed its reduction of the relevant observations and proceeded to the definition of the actual coordinate system. The intention was to choose a system which is so oriented that the direction of the pole is perpendicular to the mean plane of the Galaxy, and the zero of longitude passes through the galactic centre. These directions are known with limited accuracy and suitably rounded values for the various coordinates were chosen on the basis of considerations outlined below.

One innovation in the method of specification was made In the old system the zero of longitude was defined as one of the intersections of the galactic equator with the celestial equator for a given epoch. The new system has a zero which must be specified numerically. Various methods of specification might have been adopted, but it was decided to use the position angle at the north galactic pole with respect to the equatorial pole because this angle is one of the quantities used directly in conversion formulae from equatorial to galactic coordinates.

The system was intended to be, as far as practicable, fixed in the sky (independent of precession) and accurately reproducible. To this end the new system was defined by a geometrical definition (of infinite precision) in terms of equatorial coordinates for a given epoch. The epoch chosen was 1950.0, the standard epoch nearest to the present time. Conversions to other epochs are limited in accuracy only by the accuracy with which precession is known.

The choice of numerical values was based on the following considerations.

6.1. *The pole.*—Referring to Paper II we find that the adopted solution for the H I principal plane has a pole at the position:

$$\left.\begin{aligned}\alpha &= 12^{h}\,49^{m}\,02^{s} \pm 30^{s}\ \text{(estimated p.e.)}\\ \delta &= 27^{\circ}22'{\cdot}7 \pm 7'\ \text{(estimated p.e.)}\end{aligned}\right\}\quad (1950.0).$$

This solution is based on the region of the Galaxy within 7 kpc of the galactic centre. The radio continuum surveys are not suitable for deriving a pole, but they can be used to obtain the deviation, Δ, of the actual pole from that of the old coordinate system (Paper III). The values of Δ derived from the two most accurate continuum surveys, on the assumption that the longitude (old system) of the true pole approximates to that of the H I pole, are:

$$\Delta = 1^{\circ}{\cdot}42 \pm 0^{\circ}{\cdot}05\ \text{(p.e.)} \qquad 85\ \text{Mc/s}$$

$$\Delta = 1^{\circ}{\cdot}55 \pm 0^{\circ}{\cdot}04\ \text{(p.e.)} \qquad 1390\ \text{Mc/s}.$$

In Fig. 1 the H I pole position is plotted, with its uncertainty circle. The Δ values for the continuum are represented by short lines—corresponding to the fact that poles anywhere on these lines would yield the same Δ values. Also on the figure a slightly longer line indicates positions exactly 90° from the radio source Sagittarius A, assumed to be at the galactic centre, which has the following coordinates (Paper III)

$$\alpha = 17^h\,42^m\,37^s \qquad \delta = -28^\circ\,57' \qquad (1950.0).$$

The diagram is plotted in terms of 1950.0 equatorial coordinates, and the position of the old pole is shown for comparison.

No optical data are shown because, as concluded in Paper IV, they are based on an inadequate sample of the Galaxy, and may be influenced by observational selection. However, the radio position is concordant with the optical evidence.

Adopting the principle that the new pole is to be based primarily on the H I observations, with the radio continuum and optical results used for check purposes, a very satisfactory pole with a convenient degree of rounding in its specification is at the position

$$\alpha = 12^h\,49^m \qquad \delta = +27^\circ\cdot4 \qquad (1950.0).$$

This is shown in Fig. 1.

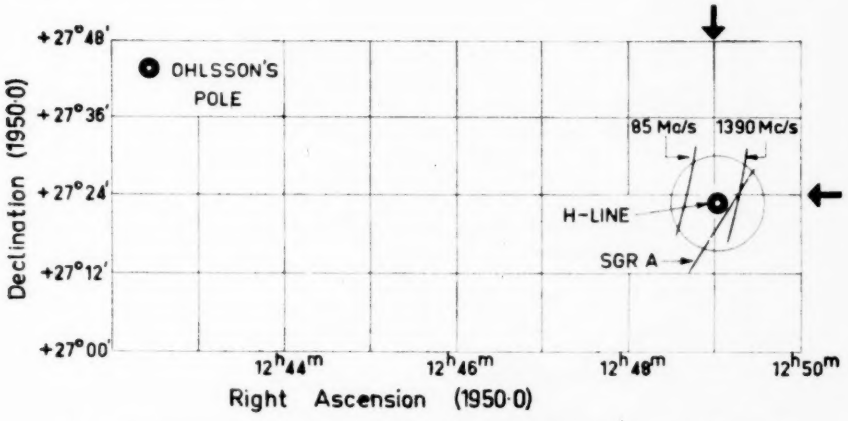


FIG. 1.—*Comparison of the position of the pole of the* H I *principal plane with various radio continuum data, and with the position of the old galactic pole. The coordinates of the new pole are indicated by heavy arrows.*

6.2. *The longitude zero.*—Adopting the above positions of Sagittarius A and the galactic pole, the position angle at the new galactic pole of the great circle passing through Sagittarius A is $\theta = 123^\circ\,00'\,30''$. This value may be conveniently rounded to $\theta = 123^\circ$.

7. *Announcement of the new system*

After reaching agreement the Commission, on 1959 March 25, communicated its definition of the new system with some explanation to the General Secretary of the I.A.U. The sections of this communication which are not already covered in the present paper are given below.

7.1. *Definition of the galactic coordinate system* (1959 *revision*):

(*a*) The new north galactic pole lies in the direction

$$\alpha = 12^{h}\,49^{m},\ \delta = +27^{\circ}{\cdot}4 \text{ (equinox 1950.0)}.$$

(*b*) The new zero of longitude is the great semi-circle originating at the new north galactic pole at the position angle

$$\theta = 123^{\circ}$$

with respect to the equatorial pole for 1950.0.

(*c*) Longitude increases from 0° to 360°. The sense is such that, on the galactic equator, increasing galactic longitude corresponds to increasing right ascension. Latitude increases from −90° through 0° to +90° at the new north galactic pole. The system is therefore similar to the Ohlsson system of coordinates.

The above quantities are to be regarded as exact so that the new galactic coordinates may be computed to any desired accuracy in terms of right ascension and declination for the equinox 1950.0.

7.2. *Other useful values* (*restricted accuracy*).—The quantities which follow have been computed from the definition but are given only to an accuracy of the nearest integer in the last digit.

(*a*) *The* 1900.0 *values corresponding to those in the definition*

New pole $\left.\begin{array}{l}\alpha = 12^{h}\,46^{m}{\cdot}6 \\ \delta = +27^{\circ}40'\end{array}\right\}$ 1900.0.

New longitude zero at position angle $\theta = 123^{\circ}\,04'$ (1900.0).

(*b*) *The old galactic coordinates of the new pole*

$$l^{I} = 347^{\circ}{\cdot}7,\quad b^{I} = +88^{\circ}{\cdot}51.$$

(*c*) *Position of the point of zero longitude and latitude* (*new system*), $l^{II} = 0$, $b^{II} = 0$
In equatorial coordinates

$$\alpha = 17^{h}\,39^{m}{\cdot}3,\quad \delta = -28^{\circ}\,54' \quad (1900.0)$$
$$\alpha = 17^{h}\,42^{m}{\cdot}4,\quad \delta = -28^{\circ}\,55' \quad (1950.0).$$

In old galactic coordinates

$$l^{I} = 327^{\circ}{\cdot}69,\quad b^{I} = -1^{\circ}{\cdot}40.$$

7.3. *Nomenclature.*—At the Moscow General Assembly it was decided that the symbols l, b should be retained for galactic longitude and latitude respectively. Commission 33b suggests that, during the transition period, the symbols l^{I}, b^{I} should be used for the old system and l^{II}, b^{II} for the new one. Apart from these superscripts, it should be made quite clear whether the galactic coordinates used in any publication are based on the old or the new system.

7.4. *Conversion tables.*—In accordance with a decision taken at the Moscow Assembly, conversion tables from equatorial into galactic coordinates and vice versa, and from the old and new galactic coordinate systems into each other, will soon be published under the supervision of Commission 33b by the Lund Observatory.

Acknowledgments.—In conclusion we wish to thank the many astronomers who assisted us in our work; in particular F. J. Kerr who has played a part in these studies equivalent to that played by the various members of the Sub-Commission.

(AB)
Kapteyn Astronomical Laboratory,
Groningen,
The Netherlands.

(CSG & JLP)
C.S.I.R.O. Radio physics Laboratory,
Sydney,
Australia.

(GW)
Sterrewacht,
Leiden,
The Netherlands.

1960 *March* 16.

References

Ashbrook, J., and Duncombe, R. L., 1952, *A. J.*, **56**, 204.
Blaauw, A., Gum, C. S., Pawsey, J. L., and Westerhout, G., 1959, *I.A.U. Information Bulletin* No. 1; also, for example, *M.N.*, **119**, 422.
Kirillova, T. S., 1955, *A. J.* (*U.S.S.R.*), **32**, 192.
Ohlsson, J., 1932, *Ann. Lund. Obs.*, **3**.
Ohlsson, J., Reiz, A., and Torgård, I., 1956, " Lund Observatory Tables for the conversion of galactic into equatorial coordinates for the Epoch 1958.0 ".
van Tulder, J. J. M., 1942, *B.A.N.*, **9**, 315.

A 21-CM DETERMINATION OF THE PRINCIPAL PLANE OF THE GALAXY

(PAPER II)

C. S. Gum, F. J. Kerr and G. Westerhout*

(Received 1960 March 21)

Summary

Observations at 21 cm have shown that the neutral hydrogen layer is exceedingly flat over the region of the Galaxy within 7 kpc from the centre. The mean plane of the hydrogen in this region must have an important dynamical significance, and may be regarded as the principal plane of the Galaxy.

The position of this plane has been determined, using data from two extensive surveys of the Milky Way by the Leiden and Sydney groups. A number of least squares solutions have been made for the position of the plane. Besides the main series of solutions incorporating all the points, various partial solutions were made for restricted regions of the Galaxy, to obtain additional information about the flatness of the hydrogen layer and the reliability of the observations. Close agreement was obtained between the best solutions, indicating that the region of the Galaxy within 7 kpc from the centre is indistinguishable from a plane.

The finally-adopted solution has the following elements and overall probable errors, expressed in terms of Ohlsson's coordinate system:

l_0^1, the longitude of the north pole of the plane, $347^\circ \pm 5^\circ$

Δ, the deviation of this pole from Ohlsson's pole, $1^\circ{\cdot}50 \pm 0^\circ{\cdot}12$

z_0, the height of the Sun above the plane, 4 pc $\pm$ 12.

1. *Introduction*

Studies of 21-cm radiation (Westerhout 1957; Kerr, Hindman and Carpenter 1957) have shown that galactic neutral hydrogen is confined to a thin layer which is very flat in the inner regions of the Galaxy and is systematically distorted in the outer parts.

The high degree of flatness of the inner region must be related to important dynamical properties of the whole Galaxy, giving special significance to the mean plane of the matter in that region (called the " principal plane of the Galaxy " by Kerr, Hindman and Carpenter (1957)). An accurate determination of its position would be expected to provide important evidence for the choice of the new galactic coordinate system.

At the present time, 21-cm observations yield the most detailed information about the three-dimensional structure in the inner region of the Galaxy, and they should therefore give the best determination of the principal plane. For definiteness, we shall refer to the plane determined from 21-cm data as the " H I principal plane." We shall not attempt to discuss here the question whether the H I solution is properly representative of all the matter in the inner region, as this subject is considered by Gum and Pawsey (Paper III). From comparisons with other radio and optical data they conclude that the available evidence is consistent with the assumption that all highly flattened subsystems of the Galaxy have a distribution which is generally similar to that of the neutral hydrogen gas.

* Deceased

The overall shape of the hydrogen layer can be seen in the relief map (Fig. 1), in which the new galactic plane has been used as the reference plane. In the relief map, as throughout this paper, the position of the neutral hydrogen layer is specified by the z-coordinate of the *maximum* hydrogen density along a line normal to the galactic plane. The available evidence indicates that the layer, whose thickness to half-density is about 200 pc, has an approximately symmetrical cross-section in the z-direction at each point.

In deriving a solution for the H I principal plane, a major problem is to ensure that the distortion in the outer region does not unduly affect the result. The form of this outer distortion (Kerr 1957; Burke 1957) may be simply expressed

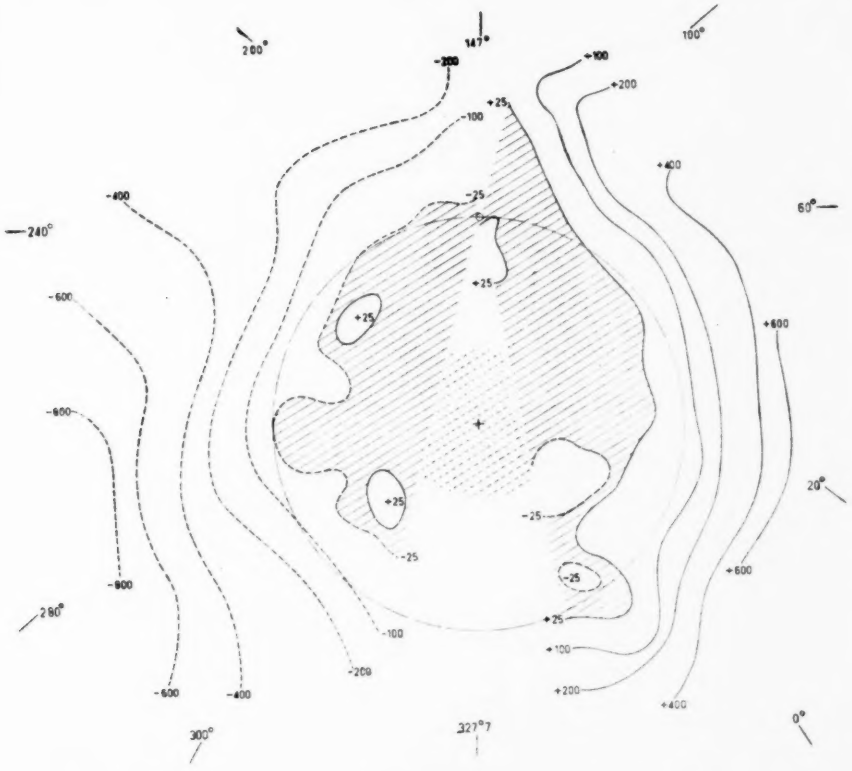


FIG. 1.—*Relief map showing the height z (pc) of the position of maximum hydrogen density above the new galactic plane. The distance scale is based on the Leiden velocity model (Westerhout* 1957; *Schmidt* 1957).

by saying that a section through the galactic centre and normal to the line joining the Sun to the centre is like a stretched-out letter " S ", with a very nearly straight central portion. (This particular section shows the distortion near its maximum.) As we are seeking to find the mean plane of the flat inner portion of the Galaxy, our problem is to decide on the extent of the inner region which

can be used for a satisfactory determination, avoiding effects due to the distorted region.

Preliminary galactic plane solutions based on 21-cm surveys have been made by Westerhout (1957) and Gum (1957) from Leiden and Sydney results respectively. In a report prepared for the (1958) General Assembly of the I.A.U. at Moscow, Gum and Kerr (1958) used data from both Leiden and Sydney, which they subdivided in various ways, in order to assess the reliability of the available observations, and the degree of flatness of the inner region of the Galaxy. The present paper is a modified version of Gum and Kerr's report, with the addition of some overall solutions based on all the 21-cm observations currently available. The preliminary solutions from the earlier report have been retained, because they contribute to the understanding and assessment of the final result.

Throughout this paper, galactic coordinates are based on the Lund pole, $\alpha = 12^h\,40^m$, $\delta = +28^\circ$ (1900) (Ohlsson 1932). They are referred to as "old" coordinates and designated by the symbols l^I and b^I.

2. *Observational material*

Extensive surveys of 21-cm radiation from the Milky Way have been made by the Leiden and Sydney groups, jointly covering a strip of sky around the galactic equator at approximately the same resolution. The observational data from the Leiden survey have been published and discussed in detail by Muller and Westerhout (1957), Westerhout (1957) and Schmidt (1957). The Sydney data have been published by Kerr, Hindman and Gum (1959), and the results have been discussed in part by Kerr, Hindman and Carpenter (1957) and Oort, Kerr and Westerhout (1958). The work described in the present paper is based on these two surveys.

We have confined our attention in the present study to observations which refer to the "inner region," where the distance from the galactic centre, R, is less than that of the Sun, R_0; the outer distortions are already obvious at $R = R_0$ and none of the solutions includes the region $R > R_0$. This restriction means that only a limited range of longitude has been considered, and only observations from one side of the velocity axis (positive velocities for northern longitudes, negative for southern).

2.1. *Leiden survey.*—Most of the Leiden observations were in the form of frequency scans across the line profile, with the aerial following particular points in the sky, but a special series of 215 drift curves was also obtained with a stationary aerial in order to study the z-distribution of the hydrogen in the inner region. The results for $R < R_0$ have been discussed by Schmidt (1957).

The drift curves used in this study are cross-sections in intensity through the Milky Way along lines of constant declination at fixed frequencies. These lines cross the old galactic equator at 5° intervals in old galactic longitude, from $l^I = 340^\circ$ to 35°. In general, two runs were made for each combination of longitude and frequency. In addition, a number of artificial drift curves were constructed for intermediate longitudes from profiles taken at two-degree intervals of latitude from $b^I = +2^\circ{\cdot}5$ to $-5^\circ{\cdot}5$.

Schmidt's results are listed in Tables 5 and 9 of his paper. The preliminary solutions in Section 5.1 below make use of the information in Table 9 only. It was realized subsequently that the data in Tables 5 and 9 are largely independent, and accordingly the results from both tables were used in the final solutions.

The observations were made at Kootwijk, with a 7·5-metre altazimuth-mounted paraboloid, whose beamwidth (between half-intensity points) was 1°·8 in azimuth and 2°·8 in elevation. The important characteristic of the aerial in studies of the z-distribution is its resolving power in galactic latitude. Over the range of longitude in which drift curves were obtained, the galactic equator is steeply inclined to the declination circles; the effective beamwidth in galactic latitude was consequently around 2°·0, since the observations were made close to the meridian.

The receiver bandwidth was 37 kc/s in frequency, equivalent to 8 km/s in velocity.

2.2. *Sydney survey.*—The Sydney observations were all obtained in the form of "galactic crossings" at fixed frequency. Over most of the range from $l^{\text{I}} = 260°$ to 5°, the records were taken as drift curves at constant declination. However, in the region $l^{\text{I}} = 260°$—280°, where declination circles are nearly parallel to the galactic equator, the aerial was driven by hand to produce a track which crossed the equator at a steeper angle.

Galactic crossings were obtained at longitude intervals of 5° over the whole range, with the addition of 6 intermediate longitudes, but in a few cases the runs were too short to give complete information on the z-distribution. At each longitude, runs were made at a sufficient number of frequencies to build up a contour diagram showing the distribution of radiation as a function of latitude and frequency (velocity). On the average, there were about 6 runs in each bandwidth.

The Sydney aerial was an 11-metre transit-mounted paraboloid, whose beamwidth was 1°·4 in both azimuth and elevation. The receiver bandwidth was 40 kc/s (8·5 km/s).

2.3. *Accuracy of position measurements.*—In considering the accuracy of the data from which the principal plane will be derived, we are concerned both with the accuracy of the measurements and with the degree of reliability of their interpretation. On the measuring side, we must study the positional accuracy of the two surveys with care, but we need not consider the intensity or frequency scales. A precise frequency measurement is the first step in a distance derivation, but adequate precision can be obtained with little difficulty.

The Kootwijk and Sydney position scales were both calibrated by means of radio observations of the Sun and a few other bright sources, and through visual checks on stars. For the 7·5-metre Kootwijk aerial, the probable error of one setting of the telescope was ±0°·07, in both azimuth and elevation (Muller and Westerhout 1957). Taking into account the uncertainty involved in measuring the position of a peak on a drift curve, the probable error in the measured galactic latitudes is estimated to be ±0°·10. For the 11-metre Sydney aerial, the corresponding figures are ±0°·05 and ±0°·07 (Kerr, Hindman and Gum 1959).

To guard against the possibility of systematic errors in either set of data, some direct comparisons have been made between the Leiden and Sydney measurements. The most precise of these comparisons was made through the discrete source, Sagittarius A (17S2A), whose emission has been measured in the continuum by both groups, with receiving arrangements essentially similar to those used in the hydrogen line work. The results are shown in Table I. The mean position derived from more recent measurements made with larger reflectors (Paper III) has been added for comparison. The Sydney and Kootwijk

values agree with one another, and with the more accurate position, to well within their probable errors.

TABLE I.

Comparison between Sydney and Leiden aerial scales through measurements on the source Sgr A.

Aerial	Beamwidth	Right ascension (1950) p.e.	Declination (1950) p.e.
Sydney-Potts Hill	1°·4	$17^h\,42^m{\cdot}7 \pm 0^m{\cdot}2$	$-29°\,01' \pm 5'$
Leiden-Kootwijk	$1°{\cdot}8 \times 2°{\cdot}8$	$17^h\,42^m{\cdot}1 \pm 0^m{\cdot}8$	$-29°\,00' \pm 15'$
Mean from larger reflectors	<0°·6	$17^h\,42^m{\cdot}6$	$-28°\,57'$

Another comparison has been made through the tangential point latitudes (see Section 3) in the longitude range in which the observations overlap. In general, the measurements were made at somewhat different longitudes, with the result that individual latitude values cannot be compared directly. Mean values of the measured latitudes over the range were therefore used, with the assumption that cosmic effects would not influence the means significantly. The Kootwijk mean was found to be higher than the Sydney mean by 0°·08, with a probable error for the difference of 0°·06. (The tangential points referred to can be located in Fig. 6, between $R=2$ and 5 kpc on the nothern side.)

This difference is hardly significant, though it is noticeable that it is in the same sense as the departure of the Kootwijk latitude for Sagittarius A from the Sydney value. We can proceed on the assumption that the two sets of data can be combined for a single treatment, but the considerations of this section will be taken into account later in assessing the probable errors of the principal plane solution.

The effects of observational uncertainties on the solutions can be estimated reasonably well, but it is difficult to make a quantitative estimate of the uncertainties arising from the analytical procedures. We will rely on comparisons between the various solutions themselves in forming an idea of the final uncertainty.

3. *Interpretation of the data*

In this section we will discuss the methods by which 21-cm observations can be interpreted, and the relative advantages of using "general" and "tangential" points for a galactic plane solution.

Each galactic crossing gives the variation of 21-cm intensity at a single frequency along a track which is steeply inclined to the galactic equator. As the variation of intensity with latitude is much greater than that with longitude, both the Sydney and Leiden groups have treated such diagonal tracks as though they were normal to the equator. Subject to this approximation, each observational result is therefore the variation of intensity with latitude at a known longitude and radial velocity. For a determination of the principal plane we wish to find the height of the neutral hydrogen layer (as represented by the point of maximum density in the z-direction) at as many points as possible in the galactic disk. Methods for making these reductions have been developed by the Leiden group (Westerhout 1957; Schmidt 1957).

The first essential is to derive a velocity-distance relation*. In principle, distances can be obtained from the measured velocities on the assumption that the major cause of line-broadening is the differential rotation of the Galaxy.

* The distance to the galactic centre has been taken as 8·2 kpc throughout this paper.

There are, however, three problems to be faced in deriving distances:

(*a*) The velocity-distance relation depends on the rotational velocity model which is chosen, and on the assumption that the gas is moving in circular orbits.

(*b*) There is in general an ambiguity in distance in the inner region, so that radiation is received at the same frequency from two points in the line of sight.

(*c*) The distribution in velocity is smeared out by the random velocities of the gas clouds.

In order to resolve the distance ambiguity, Schmidt (1957) carried out a separation of the superimposed contributions from the "near" and "far" points, on the assumption that the thickness and the z-distribution of the hydrogen remained the same all over the inner region. A substantially similar treatment has since been applied to the Sydney data. The accuracy with which the separation can be carried out is rather limited, however, because many assumptions are involved in the reduction procedure. The distances used in the present

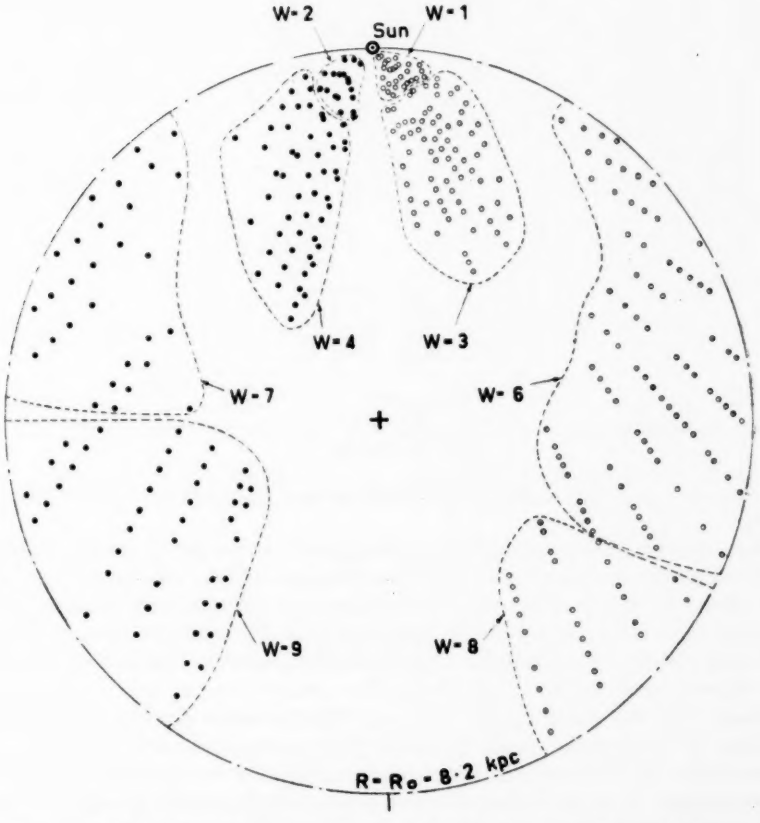


FIG. 2.—*Distribution on the galactic plane of the derived positions of general points. The weights given to the points in the preliminary least squares solutions are indicated. The data from Table 5 of Schmidt* (1957), *which were not used in the preliminary solutions, have been omitted.*

Sydney general points: ●
Leiden general points: ○

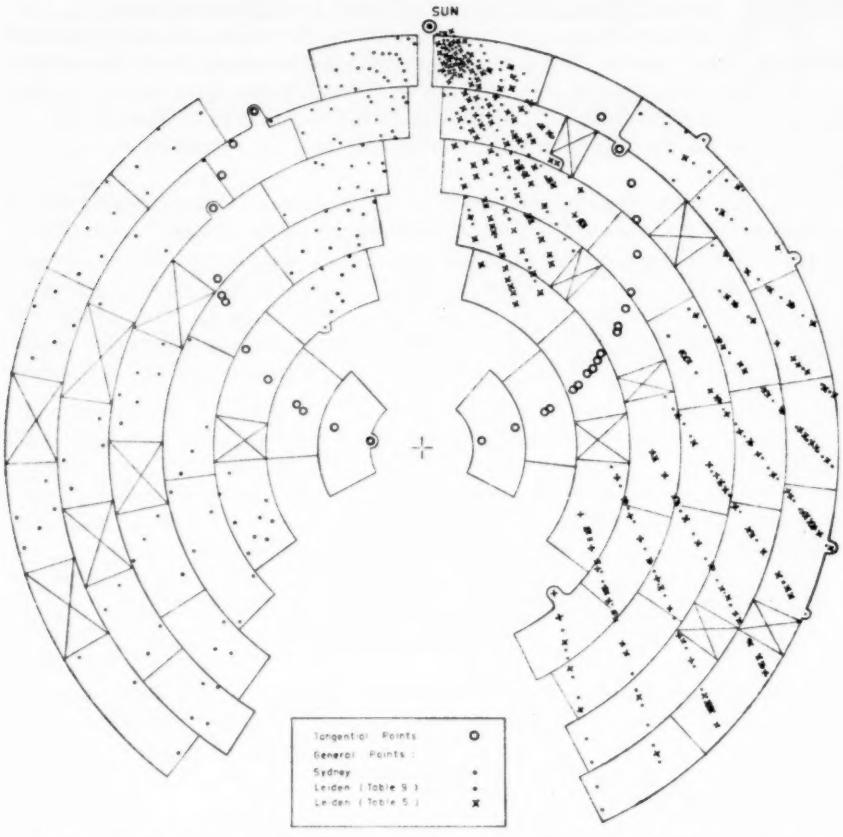


FIG. 3.—*Distribution of area elements on the galactic plane, used for normal point formation in the final solutions.*

analysis have all been based on the Leiden model (Westerhout 1957; Schmidt 1957).

We have used the phrase *general points* to refer to points which are generally distributed over the inner region, and for which the analytical separation outlined above has been used. The positions on the plane which have been derived for points of this type are shown in Figs. 2 and 3, for the points used in the preliminary and final solutions respectively. (Additional data, from Table 5 of Schmidt (1957), were used in the latter case.)

Because of the various uncertainties in the derivation of the general points, we have paid particular attention to the *tangential points*, which are the only points in the inner region where the distance ambiguity is substantially resolved. Geometrically, a tangential point is the point at which a line of sight becomes tangent to a circle about the galactic centre; here the apparent radial velocity reaches its maximum value along the line of sight. Therefore the point can be identified on the line profile, without having to consider the velocity-distance relationship. We must assume however that the gas is moving in circular orbits,

and that a substantial amount of gas is present at the geometrical tangential point. The latter assumption may not be valid at longitudes where the line of sight passes between spiral arms, but in most cases the gas which is radiating at the "maximum" velocity on a line profile is probably not far away from the true tangential point. The derived positions on the galactic plane of the 32 tangential points for which observations were available are shown in Fig. 4

In comparing solutions made from the two types of points, it should be noted that the tangential points have the advantage that their distances are better known, and the z-values do not involve the uncertainties produced by the separation process. The general points on the other hand give a better coverage of the plane, if their distances can be determined reliably.

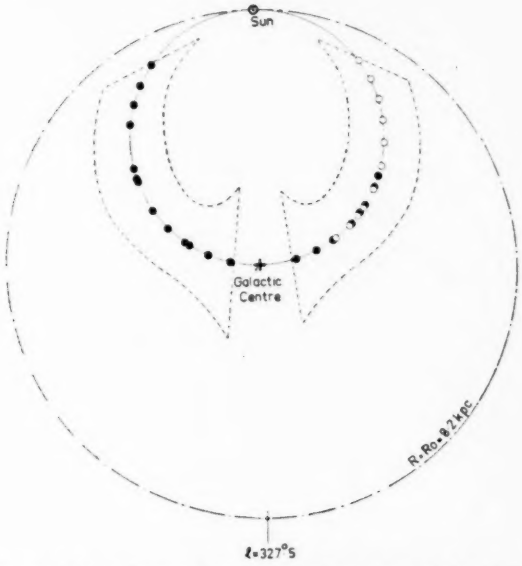


FIG. 4.—*Distribution of the 32 tangential points, and effective area of the galactic plane involved in tangential point solutions* (*Sydney* ●, *Leiden* ○).

Solutions of both types are considered in Section 5, and, in the final solutions of Section 5.2, general and tangential points are taken together. In particular, it may be pointed out that uncertainties in distance will not be very important if the Sun is found to be in or near the principal plane.

For convenience, we have been referring to "points" for which data are available, but in fact each observation refers to an appreciable area of the galactic plane. This is because the relation between the rotational velocity and the distance is blurred by the random velocities of the gas clouds, and, to a lesser extent, by the finite bandwidth of the receiver. The effective "areas of response" for some typical general and tangential points are shown in Fig. 5, for the parameters of the receivers and aerials used and the random velocity values given by Schmidt's model; the overall area which is effectively covered by the tangential points can be seen in Fig. 4. This integrating effect is to some extent an advantage, through bringing in information from a larger area of the Galaxy.

For the tangential points there is no overlap of the response areas, provided the points are spaced more than a beamwidth apart in longitude, whereas the response areas for the general points do have a considerable overlap in the radial direction. For this reason, the 32 tangential points, though few in number, refer to a substantial fraction of the whole inner region, whereas the general points do not provide as much information as their greater numbers would imply. The analytical procedures attempt to correct for this blurring effect, but in fact only a small amount of correction is possible, in view of the uncertainties in factors such as the assumed velocity dispersion law and the gas temperature.

The overall shape of the hydrogen layer can be seen from Fig. 6, which is a mass plot of the height above the new galactic plane for all the available points,

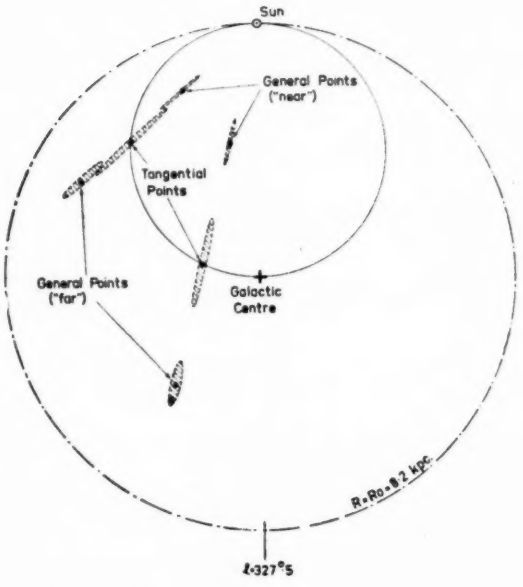


FIG. 5.—*Effective "area of response" involved in the observation of general and tangential points in the region* $R < R_0$. *The response area due to finite bandwidth, random cloud velocities, and aerial beamwidth is shown for a few typical points on two different lines of sight.*

both inside and outside the Sun's distance from the centre, except for the data from Schmidt's Table 5. Points for all galactocentric longitudes on the southern side of the Sun-centre line are plotted on the left of the diagram, points from the northern side on the right. The heights have been plotted on an expanded scale, to emphasize the deviations. It is clear that a restricted region of the Galaxy, with R less than about 7 kpc, should be used for the determination of the principal plane.

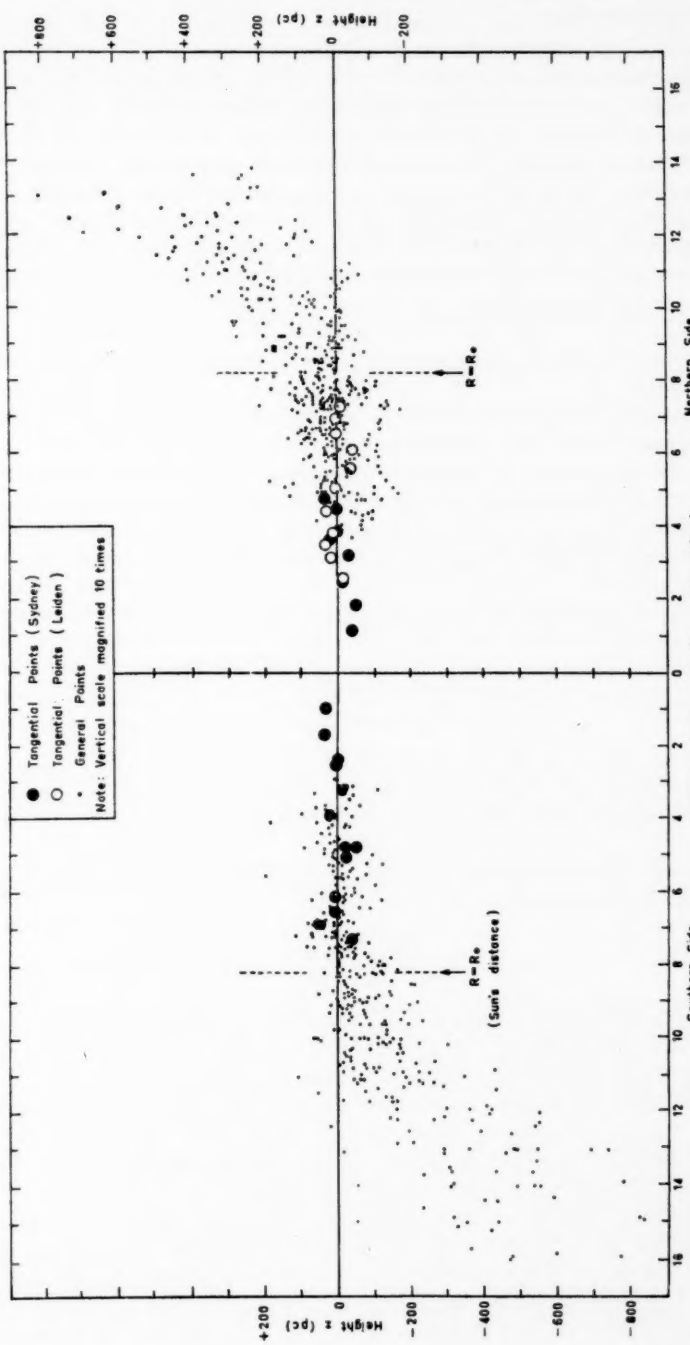


FIG. 6.—*Distribution of points of maximum hydrogen density. The heights z above the new galactic plane are plotted against the distance R from the galactic centre. Each side of the diagram refers to observations on one side of the Sun–centre line. To avoid confusion, the data from Table 5 of Schmidt* (1957) *have been omitted from the northern side.*

4. *Mathematical procedures*

4.1. *Least squares solution.*—For each solution, we have a number of points whose spatial polar coordinates are defined by r, l and b in the old galactic coordinate system, and the objective is to find by the method of least squares the mean plane which best fits the points. Since the plane we expect to find will not be far from the old galactic reference plane, it is sufficiently accurate to minimize the sum of the squares of the deviations in the direction normal to the old plane.

We shall follow the notation used by van Tulder (1942) in his determination of a galactic plane from optical data within about 3 kpc from the Sun. The height z of an observed point above the mean plane which is to be determined is given by

$$z = z_0 + r \sin b^{\mathrm{I}} \cos \Delta + r \cos b^{\mathrm{I}} \cos l^{\mathrm{I}} \sin \Delta \cos l_0^{\mathrm{I}} + r \cos b^{\mathrm{I}} \sin l^{\mathrm{I}} \sin \Delta \sin l_0^{\mathrm{I}}$$

where the orientation is such that positive z values are on the northern side of the plane, and where

z_0 is the height of the Sun above the mean plane to be determined,
r is the distance of an observed point from the Sun,
l^{I}, b^{I} are the old galactic coordinates of the point,
l_0^{I} is the old longitude of the north pole of the mean plane,
and Δ is the deviation of this pole from the old pole.

Since the angles Δ and b^{I} are both small, we have, with ample accuracy for our purpose,

$$z = z_0 + r \sin b^{\mathrm{I}} + (r \cos l^{\mathrm{I}})X + (r \sin l^{\mathrm{I}})Y \qquad (1)$$

where

$$X = \sin \Delta \cos l_0^{\mathrm{I}}$$
$$Y = \sin \Delta \sin l_0^{\mathrm{I}}. \qquad (2)$$

For n points, each defined by a set of values (r,l^{I} and b^{I}), a mean plane is to be determined such that $\Sigma w z^2$ is a minimum. Each point yields an equation of condition of the form of equation (1), whose weight w is that assigned to the point.

The internal error in the determination of the quantities X, Y and z_0 may be given by an expression of the form

$$\epsilon = \pm \sqrt{\frac{A_{pp} \Sigma w z^2}{D(n-m)}}$$

where m is the number of unknowns and A_{pp} and D are co-factors and determinant of the matrix of coefficients of the n equations of condition. In this paper we shall quote probable errors, which are 0·67 times the error ϵ.

When the errors in X and Y have been derived they must be converted to errors in l_0^{I} and Δ by means of equations (2). Since these equations involve non-linear functions, the resulting errors will be slightly asymmetrical. The effect is only noticeable when the errors are fairly large.

4.2. *Weighting.*—In order to find a properly representative plane for the inner region, the points must be appropriately weighted; otherwise the regions of the Galaxy near the Sun where the observed points are more closely packed would have too strong an influence on the solution. Ideally it would be desirable to reach the equivalent of an even distribution of points over the whole of the region of the Galaxy concerned in a particular solution. In practice, there are certain

areas of the Galaxy where no observed points exist, such as the directions near that of the galactic centre, where the 21-cm analysis breaks down (see Fig. 2). One must therefore be content with the equivalent of an even distribution of points over the areas within which observed points do occur.

In all the solutions, weights have been assigned on the basis that the weight of the points in a given "locality" on the plane should be inversely proportional to the "local density" of the points per unit area of the plane. Different methods of treatment were however used for the preliminary and final solutions.

In the preliminary solutions, the general points were weighted in groups by finding the average density of points per kpc^2 for eight relatively large areas into which the plane was divided. These areas were chosen so that the fluctuations of local density within each area would not be too large. The tangential points received weights which were inversely proportional to their linear density along the "midpoint circle" (see Fig. 4). For both general and tangential points, each individual point led to a separate equation of condition with appropriate weight.

In the final solutions, the amount of labour in the least squares solutions was reduced by forming normal points for the effectively equal area elements of Fig. 3. These area elements are in general bounded by lines of constant R (with integral values in kpc), and extend over ranges of galactocentric longitude which were chosen to make all the areas equal. However, variations of up to 50 per cent in element area were allowed to occur so that no element would contain too small a number of points. On the average there are 13 general points or 2 tangential points per element. This is in keeping with the result that the tangential points show the smaller scatter, so that they should receive greater weight than the general points. (The r.m.s. deviation of an individual tangential point from the mean plane is 25 pc, and that of a general point 70 pc.)

We may remark that it is neither necessary nor possible to achieve a high degree of precision in the weighting for a somewhat irregular distribution of points. To confirm that a satisfactory weighting system had been arrived at, some check solutions were made, using first the normal point method and then the individual point method for the same sets of data. The corresponding solutions agreed with each other to well within their probable errors, indicating that the two weighting systems are effectively similar.

5. *Solutions for the position of the* H I *principal plane*

Before discussing the "final" solutions, we will first consider the large variety of preliminary solutions which were made before the 1958 I.A.U. General Assembly. Intercomparisons between these preliminary solutions provide valuable information about the reliability of the observations and the characteristics of the galactic hydrogen layer.

5.1. *Preliminary solutions.*—The preliminary solutions numbered 25 in all; separate solutions were made for tangential and general points, in each case for various groupings of the data corresponding to different regions of the Galaxy.

For the tangential point solutions, the Leiden and Sydney observations were taken together. The general point solutions are from Leiden and Sydney data taken separately as well as in combination. The Leiden general points refer to regions of the Galaxy on the northern side of the line through the Sun and the galactic centre, while the Sydney general points refer to the southern side (see Fig. 2).

Each group of solutions consists of a series which refers to data within different distances $R < R'$ from the galactic centre. These partial solutions were intended to provide information on the value of R at which the outer distortion sets in. Another feature is that each group of solutions applies to a different part of the inner region of the Galaxy, and hence comparisons between the groups give extra evidence on the flatness of the inner region. Further, since the reductions for the tangential and general points are based on rather different assumptions and interpretations, we have some check on possible effects arising from the assumptions involved.

The results of all the preliminary solutions are listed in Table II, which also gives information on the region concerned in each solution, and the type of data used. The elements of the various solutions are also shown diagramatically in Fig. 7. For each solution, the values of the elements l_0^{I}, Δ and z_0, and their errors, are found vertically below the identification number at the top of the diagram, while the limit R' for the solution is found below*.

The errors quoted in the table and in Fig. 7 are the *internal* errors for each least squares solution, ignoring for the time being the fact that the z-values and the distances may contain systematic errors, arising in the observations or the analysis.

Several conclusions can be drawn from comparisons between the various preliminary solutions:

(*a*) The solutions with limiting R-values of 7·5, 7·0 and 6·5 kpc show a high degree of consistency. This is especially so for the solutions based on the combined general points (Nos. 22, 23 and 24). It may be recalled that a series of solutions with varying R' was made in order to obtain a basis for selecting the region of the Galaxy to be used for the final determination. We hoped to find a region which was small enough to exclude the outer systematic distortion, but large enough to contain a sufficient number of observed points. From a study of these solutions, and of the relief map, we decided that the best compromise would be to take the region inside $R = 7{\cdot}5$ kpc.

We may remark here that there is a suggestion of a small amount of "waviness" of the hydrogen layer in the inner region, but the z-amplitude could only be about 20 pc or so. Because this is small compared with the 800 pc outer distortion, such a feature would not seriously affect the concept of the principal plane of the Galaxy.

(*b*) Provided the region is restricted to $R' = 7{\cdot}5$ kpc or less, solutions based on Leiden data alone agree well with those based only on Sydney data. This result provides additional evidence that the Leiden and Sydney position scales are in good agreement and also that the hydrogen layer lies approximately in a single plane over the whole inner region of the Galaxy.

(*c*) Again with the proviso that $R < 7{\cdot}5$ kpc, solutions based on general and tangential points agree well with each other. Consequently, we can combine both sets of points in our final solutions, in order to use the largest possible body of information.

5.2. *Final solutions.*—The group of final solutions had the following characteristics:

(*a*) They were restricted to R' values of 8·0, 7·0, 6·0, 5·0 kpc.

* The identification numbers of the solutions are different from those used in the earlier report Gum and Kerr 1958).

TABLE II

Results of preliminary solutions

	Soln. No.	R' (kpc)	No. of points	l_0^{I} p.e.	Δ p.e.	z_0 (pc) p.e.
General points, northern side only (Leiden data)	1	5·0	30	317° $^{+20^\circ}_{-15}$	$2^\circ{\cdot}0$ $^{+1^\circ{\cdot}5}_{-0{\cdot}6}$	+12 ±29
	2	5·5	52	312 ±10	2·2 $^{+1\cdot0}_{-0\cdot5}$	− 5 ±22
	3	6·0	74	333 ±14	1·67 ±0·34	+ 2 ±20
	4	6·5	96	338 ± 9	1·63 ±0·20	− 2 ±14
	5	7·0	128	343 ± 7	1·57 ±0·13	− 3 ±11
	6	7·5	165	334 ± 5	1·61 ±0·14	0 ±10
	7	8·2	209	329 ± 5	1·54 ±0·11	+13 ± 9
General points, southern side only (Sydney data)	8	5·0	38	348 $^{+22}_{-26}$	1·34 $^{+1\cdot1}_{-0\cdot4}$	+10 ±23
	9	5·5	49	337 $^{+13}_{-18}$	1·30 $^{+0\cdot9}_{-0\cdot4}$	+ 2 ±18
	10	6·0	63	314 ±16	1·05 $^{+1\cdot5}_{-0\cdot4}$	+ 8 ±15
	11	6·5	78	322 ±10	1·26 $^{+0\cdot6}_{-0\cdot3}$	−10 ±13
	12	7·0	97	347 ± 5	1·57 ±0·20	− 7 ±11
	13	7·5	120	352 ± 4	1·66 ±0·11	0 ± 9
	14	8·2	142	336 ± 5	1·35 ±0·15	− 4 ± 9
Tangential points, northern and southern observations (Leiden and Sydney data)	15	5·0	16	349 ± 5	1·56 ±0·25	0 ±32
	16	6·3	26	351 ± 5	1·42 ±0·16	+23 ±17
	17	$3{\cdot}2<R<7{\cdot}5$	23	346 ± 4	1·55 ±0·15	+ 1 ±10
	18	7·5	32	348 ± 3	1·59 ±0·11	− 1 ±10
General points, northern and southern observations (Leiden and Sydney data)	19	5·0	68	358·6 ±3·7	1·79 ±0·11	− 1 ±16
	20	5·5	101	350·2 ±3·1	1·76 ±0·10	−18 ±13
	21	6·0	137	349·7 ±2·8	1·59 ±0·08	+ 1 ±11
	22	6·5	174	345·7 ±2·4	1·59 ±0·07	− 4 ± 9
	23	7·0	225	346·8 ±2·0	1·57 ±0·06	− 6 ± 7
	24	7·5	285	346·1 ±1·8	1·57 ±0·05	−10 ± 6
	25	8·2	351	341·7 ±1·7	1·48 ±0·05	0 ± 5

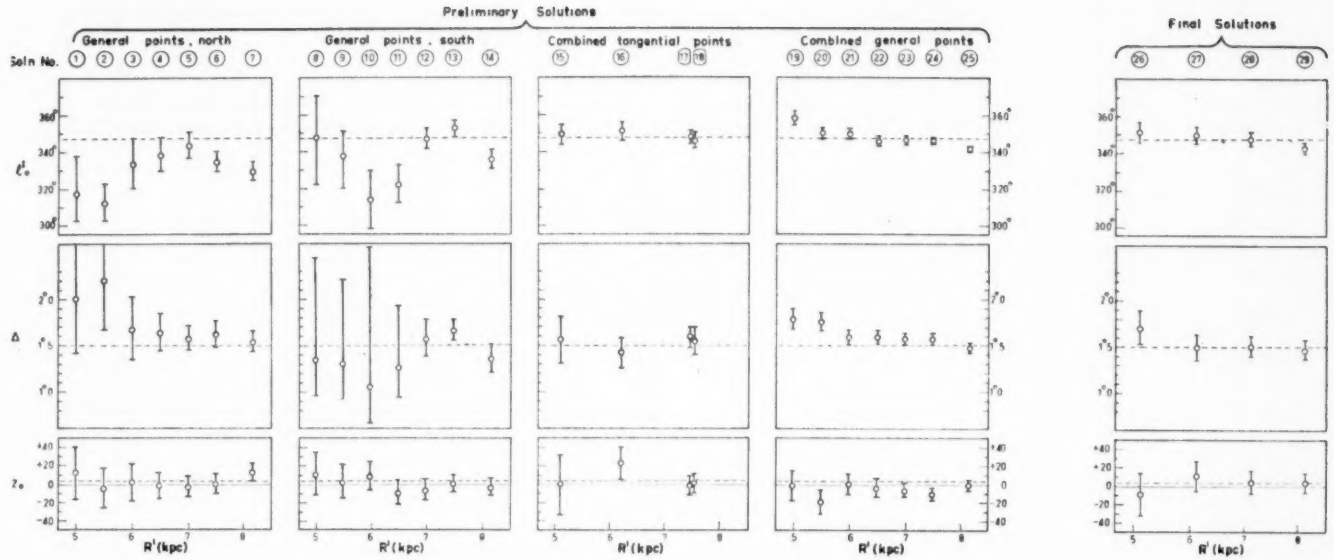


FIG. 7.—*Results of the 25 preliminary and 4 final solutions. For a particular solution, the values of the three elements l_0^I, Δ and z_0, together with their probable errors, are plotted vertically below the identification number at the top of the diagram. The relevant upper limit R' for the solution is given at the bottom. The dashed lines indicate the elements of the finally-adopted solution.*

TABLE III

Normal points used in final solutions for H I *principal plane*

No.	Heliocentric coordinates		Galactocentric coordinates		Heights	
	r (kpc)	l^{II}	R (kpc)	L^{II}	z^{I} (pc)	z^{II} (pc)
1	8·03	338°·5	1·54	102°·1	−260	−52
2	7·78	345·9	2·57	108·3	−226	−22
3	9·95	347·7	3·59	71·3	−300	−39
4	7·37	353·4	3·55	116·1	−189	+ 2
5	5·26	347·1	3·68	161·6	−155	−17
6	11·97	342·6	4·55	42·4	−356	−44
7	11·00	348·2	4·38	61·3	−353	−65
8	9·40	356·3	4·50	90·0	−234	+ 9
9	6·79	1·5	4·57	124·1	−154	+18
10	4·27	350·4	4·57	158·8	−128	−17
11	12·87	343·9	5·49	40·8	−275	+61
12	11·44	353·3	5·38	66·8	−347	−50
13	9·99	1·2	5·52	86·5	−243	+11
14	8·21	7·1	5·53	109·6	−207	− 5
15	6·19	8·2	5·33	131·0	−168	−17
16	4·14	4·8	5·50	153·0	− 94	+ 9
17	3·12	352·6	5·53	166·3	− 80	+ 1
18	13·48	347·1	6·35	44·9	−293	+60
19	12·00	357·9	6·42	70·3	−262	+46
20	10·13	7·7	6·53	85·9	−269	−21
21	8·12	14·3	6·45	114·5	−210	−21
22	5·17	18·2	6·33	140·9	−134	−18
23	2·30	3·6	6·48	168·0	− 39	+19
24	14·90	343·9	7·39	34·1	−441	−52
25	13·93	353·2	7·43	53·7	−298	+64
26	12·25	4·3	7·49	77·4	−231	+75
27	10·76	12·5	7·60	94·4	−188	+67
28	9·16	18·5	7·49	108·7	−130	+75
29	7·15	25·7	7·50	126·0	−144	+ 3
30	5·42	32·4	7·66	140·2	− 99	+ 1
31	4·06	27·7	7·10	150·3	− 78	+ 3
32	1·01	13·1	7·52	174·5	− 8	+16
33	8·08	318·4	1·33	260·3	−158	+27
34	7·83	310·3	2·44	253·0	−170	− 6
35	5·44	308·9	3·51	209·9	−114	− 3
36	7·40	302·3	3·52	244·4	−140	− 3
37	10·12	308·9	3·54	292·9	−219	−12
38	4·41	304·6	4·49	202·6	− 70	+15

Table III (*cont.*)

No.	Heliocentric coordinates		Galactocentric coordinates		Heights	
	r (kpc)	l^{II}	R (kpc)	L^{II}	z^{I} (pc)	z^{II} (pc)
39	6·66	291°·8	4·81	234°·3	−139	−40
40	8·99	297·9	4·48	265·1	−158	− 4
41	10·99	306·1	4·61	300·1	−177	+39
42	3·35	299·5	5·49	196·8	− 75	−16
43	5·92	285·0	5·56	226·2	− 80	− 8
44	8·98	290·2	5·57	258·9	−160	−33
45	11·44	301·0	5·54	291·8	−249	−42
46	12·58	310·2	5·36	315·1	−215	+48
47	2·26	291·9	6·50	191·7	− 28	+ 6
48	4·36	270·2	6·91	212·1	− 21	+ 5
49	6·79	276·2	6·65	233·0	− 20	+38
50	10·69	290·4	6·48	272·7	−143	+10
51	12·83	302·2	6·47	301·5	−252	−15
52	13·79	310·4	6·44	320·5	−344	−56
53	1·24	277·0	7·48	187·4	− 17	− 6
54	5·70	264·1	7·72	221·9	− 42	−24
55	7·70	270·6	7·61	238·1	− 99	−42
56	9·36	277·3	7·55	252·8	−147	−37
57	12·18	290·4	7·53	281·3	−162	+12
58	14·24	304·8	7·40	311·6	−332	−57

(*b*) All available general and tangential points were included in each solution, the overall total being 463 points. The northern-side points given in Table 5 of Schmidt (1957) were added to the Leiden and Sydney data which had been used in the preliminary solutions.

(*c*) The region was divided into 58 approximately equal areas (see Fig. 3), for which normal points were formed.

The normal point data are presented in Table III. In this Table the positions of points in the galactic plane are specified in both heliocentric and galactocentric polar coordinates. In the former case the angular coordinate is simply the old galactic longitude; in the latter the origin of coordinates is translated without rotation).

The four solutions in this group are given in Table IV, and also shown diagramatically in Fig. 7.

TABLE IV

Results of the four "final" solutions

R' (kpc)	l_0^{I}	p.e.	Δ	p.e.	z_0 (pc)	p.e.
5·0	350°·6	±5·6	1°·70	±0·18	− 9	±23
6·0	348°·9	±4·6	1°·49	±0·13	+11	±16
7·0	347°·2	±3·5	1°·50	±0·11	+ 4	±12
8·0	341°·8	±3·2	1°·46	±0·10	+ 3	±10

Taking the solution for $R' = 7{\cdot}0$ kpc, the finally-adopted values for the elements are:

$$l_0^{I} = 347°, \quad \Delta = 1°{\cdot}50, \quad \text{and} \quad z_0 = +4 \text{ pc}.$$

These values have been included in Fig. 7 as dashed horizontal lines, so that the final solution may be compared with the whole range of individual solutions.

The errors quoted in this section have been the internal errors obtained in the least squares solutions; these are internal errors only and are related to the random scatter of the observed points about the mean plane determined in each case. These errors take no account of systematic effects, which must be considered separately. For example, a systematic difference of 0°·05 between the Sydney and Leiden declination scales would produce a change of about 1°·5 in $l_{\odot}^{\mathrm{I}}$. If there were an absolute systematic error of 0°·05 in the same sense in both declination systems, there would be a resultant error of about 0°·05 in Δ.

Systematic errors can also arise from the analysis procedures, or from large-scale departures of the hydrogen layer from a plane. We cannot make an estimate of these uncertainties, but comparisons between the various solutions imply that they cannot be very large. In particular, there is close agreement between the solutions from general and tangential points, which are based on very different assumptions. In assessing the overall uncertainty, we first estimated an internal probable error for the whole group of preliminary and final solutions. This error is less than the errors for individual solutions, because of the consistency of the group; the various solutions were only partially independent, but their consistency indicates that the final result is not very dependent on the precise method of treatment of the data. Finally, an allowance was made for possible systematic errors.

Our finally-adopted solution for the principal plane has the following elements and estimated overall probable errors:

$$l_0^{\mathrm{I}} = 347^\circ \pm 5^\circ \text{ (estimated p.e.)}$$
$$\Delta = 1^\circ{\cdot}50 \pm 0^\circ{\cdot}12 \text{ (estimated p.e.)}$$
$$z_0 = 4 \text{ pc} \pm 12 \text{ (estimated p.e.)}$$

The position of the corresponding pole in equatorial coordinates is given by:

α	δ	
$12^{\mathrm{h}}\,46^{\mathrm{m}}\,36^{\mathrm{s}}$	$+27^\circ\,39'{\cdot}0$	(1900.0)
$12^{\mathrm{h}}\,49^{\mathrm{m}}\,02^{\mathrm{s}}$	$+27^\circ\,22'{\cdot}7$	(1950.0).

(CSG & FJK)
C.S.I.R.O. Radiophysics Laboratory,
Sydney,
Australia.

(GW)
Sterrewacht,
Leiden,
The Netherlands.

1960 *March* 16.

References

Burke, B. F., 1957, *A.J.*, **62**, 90.
Gum, C. S., 1957, Mt. Stromlo Pre-Conference on the Coordination of Galactic Research, 16.
Gum, C. S., and Kerr, F. J., 1958, C.S.I.R.O. Radiophysics Lab. Rept. 138.
Kerr, F. J., 1957, *A.J.*, **62**, 93.
Kerr, F. J., Hindman, J. V. and Carpenter, M.S., 1957, *Nature*, **180**, 677.
Kerr, F. J., Hindman, J. V., and Gum, C. S., 1959, *Aust. J. Phys.*, **12**, 270.
Muller, C. A., and Westerhout, G., 1957, *B.A.N.*, **13**, 151.
Ohlsson, J., 1932, *Lund. Ann.* No. 3.
Oort, J. H., Kerr, F. J., and Westerhout, G., 1958, *M.N.*, **118**, 379.
Schmidt, M., 1957, B.A.N., **13**, 247.
Van Tulder, J. J. M., 1942, *B.A.N.*, **9**, 315.
Westerhout, G., 1957, *B.A.N.*, **13**, 201.

RADIO DATA RELEVANT TO THE CHOICE OF A GALACTIC COORDINATE SYSTEM*

(PAPER III)

C. S. Gum† and J. L. Pawsey

(Received 1960 March 21)

Summary

Radio observations relating to the determination of the directions towards the galactic pole and centre are collected and studied. The most important evidence on the direction of the pole is based on 21-cm hydrogen line observations; the derived direction is confirmed by observations in the radio continuum. The direction towards the galactic centre is deduced from various continuum and hydrogen line observations, the most precise information being derived from the position of the radio source Sagittarius A which is assumed to be associated with the galactic nucleus and at the galactic centre.

The outstanding physical result of this study is the extreme flatness of the disk of neutral hydrogen in its central parts; a new point is the large-scale association between the sources of the radio continuum and the 21-cm line radiation. The significance of the results is discussed.

1. *Introduction*

Historically, the first radio evidence on the position of the galactic plane came from surveys at various frequencies in the radio continuum. There was a general indication that, in the vicinity of the galactic centre, the line of maximum brightness, the "ridge-line", lay a degree or so south of the (old) galactic equator, but the results were not easy to interpret. In fact, two groups of workers who used continuum results to estimate the position of the galactic plane obtained results differing significantly from what we now believe to be correct, owing to quite natural errors in interpretation.

It was the discovery and utilization of the 21-cm hydrogen line which clarified the situation. Using this line it became possible to estimate distances and hence to derive a three-dimensional distribution of the neutral hydrogen gas which is concentrated in the galactic disk. This information led to the recognition of the extreme flatness of the hydrogen layer in the inner parts of the Galaxy and its systematic distortion in the outer parts.

By contrast, observations in the radio continuum yield only integrated effects along the line of sight. At the lower frequencies the main contribution to the sky brightness will be from regions near the Sun; at higher frequencies the transparency will be greater and the whole Galaxy will contribute, but as yet there is no way to derive a three-dimensional distribution of the sources in the disk from the continuum results alone. The improvement in interpretation

* This paper is essentially a condensation of a report (Radiophysics Laboratory Report RPL 137) with the same title prepared for I.A.U. Sub-Commission 33b. It differs from the report in two respects: (1) the omission of certain material now covered by other papers of the present series, and (2) the inclusion of some results obtained subsequent to June 1958, which have been generously communicated to us prior to publication.

† Deceased.

using the H-line data is analogous to the improvement realized optically by three-dimensional studies, such as those of van Tulder, over earlier studies using star counts (e.g. those of Herschel) or photographic isophotes (e.g. those of Barnard).

Despite the additional power of the H-line observations, we have studied the continuum results very carefully for two main reasons. Firstly the H-line results refer to a single sub-system in the Galaxy, neutral interstellar hydrogen. It would be expected that the hydrogen gas would be symmetrically disposed with regard to the mean plane of the whole Galaxy, but similar results with respect to any other sub-system would supply useful confirmatory evidence. We have been able to show that the continuum results do in fact provide such evidence. While there is still some uncertainty as to the physical nature of the sources responsible for the continuum emission—or rather for the non-thermal part of the emission—this feature is of no great concern for our present purposes.

Secondly the H-line results considered here are restricted to only two series of observations, those from Leiden and Sydney. The use of continuum results from a much larger number of observers for check purposes therefore reduces the possibility of systematic observational error.

In comparison with optical data, the outstanding feature of the radio data is the freedom from interstellar obscuration. For the H-line and the shorter wavelengths of the continuum the whole Galaxy is accessible to observation. This is the feature which finally led to the selection of the radio data as the basis for the choice of the new galactic plane. The greater angular accuracy of optical observations is more than offset by the much better sample of the Galaxy accessible to radio observation.

In the case of the location of the galactic centre, radio and optical studies led to accuracies of a fraction of a degree. This uncertainty is reduced to a few minutes of arc when the radio source Sagittarius A, which is considered to be at the galactic centre, is used for locating the centre.

2. *Evidence relevant to the choice of a galactic pole*

The Galaxy is a rotating flattened conglomeration of matter in which certain types of objects are concentrated towards the equatorial plane in differing degrees. In principle one might choose the direction to the pole of a galactic coordinate system to be (1) parallel to the axis of rotation, or (2) normal to the mean plane of the whole Galaxy, or (3) normal to the plane defined by the strongly concentrated objects. In practice the last can be much more accurately located and has been used exclusively. The sources of the radio continuum and the 21-cm line are both highly concentrated, and the disk defined by the concentration is remarkably flat in the inner parts. Following the concept of Kerr, Hindman and Carpenter (1957), we shall designate as the "principal plane" of the Galaxy the mean plane in the inner region for all matter showing high galactic concentration. At the present time it is possible to determine *directly* the position of the principal plane only from the distribution of the H I clouds. The plane defined by those clouds we designate the "H I principal plane".

This section presents the 21-cm and continuum evidence on the location of the principal plane and discusses the way in which the respective sources are distributed.

Throughout this paper, galactic coordinates are based on the Lund pole, $\alpha = 12^h 40^m$, $\delta = +28°$ (1900) (Ohlsson 1932). They are referred to as "old" coordinates and designated by the symbols l^I and b^I.

2.1. *The 21-cm evidence.*—A full discussion of the 21-cm evidence is given in Paper II. The best estimate of the position of the H I principal plane, expressed in terms of old galactic coordinates, is:

$$l_0^{\mathrm{I}} = 347^\circ \pm 5^\circ \text{ (estimated p.e.)}$$
$$\Delta = 1^\circ{\cdot}50 \pm 0^\circ{\cdot}12 \text{ (estimated p.e.)}$$
$$z_0 = +4 \text{ pc} \pm 12 \text{ (estimated p.e.)}$$

where

l_0^{I} is the longitude of the north pole of the H I plane,
Δ the deviation of this pole from the old one, and
z_0 the height of the Sun above the H I plane (positive for northerly displacements).

We shall later discuss the significance of the extreme flatness of the central regions of the H I disk, and the distortions in the outer parts, best presented in the relief map on p. 133 (Fig. 1 in Paper II).

2.2. *The continuum evidence.*—A number of radio continuum surveys covering various parts of the sky are available on frequencies ranging from 40 to 1400 Mc/s. These surveys give a two-dimensional picture of the distribution of radio emission over the sky, each showing a ridge of high intensity broadly following the Milky Way. The position of the "ridge-line" of maximum intensity can be derived in each case.

The surveys considered here are listed in Table I.

TABLE I

Surveys in the radio continuum

Wide-Beam Surveys:	Frequency Mc/s	Beamwidth
Reber (1948)	480	4°
Bolton and Westfold (1950)	100	17°
Allen and Gum (1950)	200	25°
Seeger and Williamson (1951)	205	15°
Ko and Kraus (1957)	250	8° × 1°
Baldwin (1955)	81·5	17° × 2°
Denisse, Leroux and Steinberg (1955)	910	≈3° × 4°
Dröge and Priester (1956)	200	16°·5
Blythe (1957)	38	2°·2 × (2°·3 to 7°·4)
Narrow-Beam Surveys:		
Hanbury Brown and Hazard (1953)	158·5	2°
McGee, Slee and Stanley (1955)	400	2°
Piddington and Trent (1956)	600	3°·3
Hill, Slee and Mills (1958)	85·5	0°·8
Westerhout (1958)	1390	0°·57

Certain longitude ranges in some surveys have been neglected. These are usually cases where the scale of published diagrams is rather small or contours are so widely spaced that reading errors are serious. For each survey in the table the ridge-line of maximum radiation has been drawn, omitting sections where it is obviously influenced by discrete sources.

The surveys have been separated into two groups—one, which we call "wide-beam surveys", includes all those with aerial beamwidths to half power

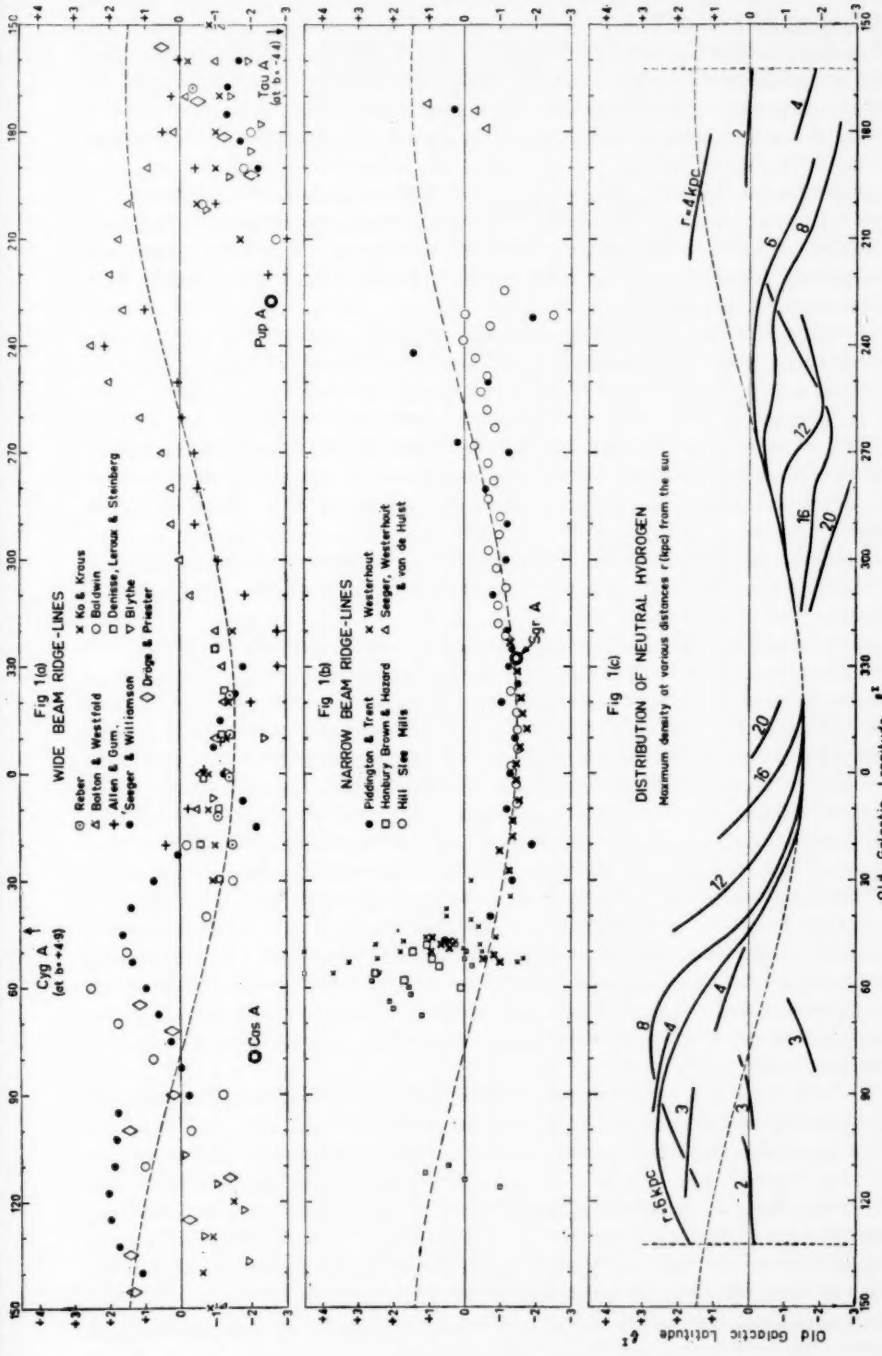


FIG. 1.—*(a) and (b): Positions in the sky of radio continuum "ridge-lines"; (c): corresponding positions of regions of maximum density of neutral hydrogen. The sine curve common to the three figures gives the position of the new galactic equator.* (Old galactic coordinates.)

greater than $3\frac{1}{2}$ degrees, the other "narrow-beam surveys" with beamwidths less than $3\frac{1}{2}$ degrees. The ridge-line data for the wide-beam group are shown in Fig. 1(*a*), and for the narrow-beam group in Fig. 1(*b*)*. For the narrow-beam ridge-lines, effectively independent points are available at intervals of about one degree. In order to avoid the confusion produced by such a large number of points, we have plotted in Fig. 1(*b*) mean points over an interval of 5 degrees of longitude throughout the region where the ridge-line is well defined. For longitudes where the ridge-line is not so well defined, and much fine structure is apparent, the positions of individual maxima, as well as the position of the centroids of sections normal to the galactic equator, are shown. A limited number of Cambridge "ridge-line" positions, kindly provided by F. G. Smith, were consistent with those shown.

A striking feature of both Figs. 1(*a*) and (*b*) is the small scatter in the galactic latitude of the ridge-lines over a range of about 120 degrees of longitude around the galactic centre. It is clear that the ridge-line is intrinsically less well defined for longitudes further from the galactic centre than 60 degrees. This is presumably due in part to the fact that in directions away from the centre the radiation originates in fewer and closer sources, but systematic distortions of the galactic "plane" probably account for much of this effect (see next section).

A sine curve common to Figs. 1(*a*), (*b*) and (*c*) represents the adopted solution for the "principal plane" of neutral hydrogen gas in the inner regions of the Galaxy.

2.3. *Comparison of hydrogen line and radio continuum data.*—Within about 60 degrees on either side of the galactic centre, the positions of the ridge-lines of maximum intensity of the continuum surveys agree very well with the position of the H I principal plane (Fig. 1(*b*)). This indicates that in the inner regions of the Galaxy the "disk" component of the continuum radiation is concentrated towards the same plane of symmetry as the neutral hydrogen gas. (The "corona" component in the continuum is much broader in galactic latitude, and, though it appears to be concentrated to the same plane, can only be located very roughly). For longitudes further than 60 degrees from the galactic centre, the continuum ridge-lines depart in a fairly systematic way from the position of the principal plane. Since for these longitudes the hydrogen layer also shows systematic distortions, we are led to pose the following question: Supposing that the distribution in space of the sources of the "disk component" of the continuum radiation were generally similar to that of the neutral hydrogen, what would be the general appearance of the continuum surveys, in particular the positions of their ridge-lines?

To help towards a study of this question, a three-dimensional representation of the position of the hydrogen layer has been plotted in Fig. 1(*c*). In this figure, which gives the same information as the more conventional relief map (see Fig. 1 in Paper II, p. 133), the principal plane is represented by the dashed sine curve while the position of the distorted parts of the hydrogen layer at different distances r from the Sun are shown by various lines. Thus the figure gives the "sky positions" of the H I layer for various distances from the Sun. For distances from the Sun $r < 4$ kpc, the relief map does not provide sufficiently

* The narrow-beam data of McGee, Slee and Stanley (1955) are not shown because their data appear to contain a systematic error in galactic latitude.

detailed information owing to the considerable amount of smoothing necessary in order to produce continuous contours. The use of additional unsmoothed data, mainly from Westerhout's Fig. 9 (1957), accounts for the $r=2$, $r=3$ and $r=4$ kpc data appearing as disconnected lines in Fig. 1(*c*). It is important to include these data, particularly in the region of the anticentre, where the relative nearness to the Sun results in a considerable scatter in galactic latitude not shown by the relief map.

Figs. 1(*a*) and 1(*b*) may now be compared with what would result if Fig. 1(*c*), besides giving the three-dimensional distribution of the H I gas, also represented a model of the distribution of the radio continuum sources. Without additional assumptions it is, of course, impossible to compute the variation of the expected brightness temperature as a function of galactic latitude, but a qualitative comparison is still possible. If our supposition is correct, then the continuum ridge-lines (Figs. 1(*a*) and 1(*b*)), being the maxima of the integrated radiation along the line of sight, should fall somewhere inside the latitude range of the sky positions of the H I layer in Fig. 1(*c*).

It will be seen that the narrow-beam ridge-lines (Fig. 1(*b*)) satisfy this test. In the vicinity of $l^{\text{I}}=230^\circ$ to 270° they fall consistently below the principal plane, and from $l^{\text{I}}=20^\circ$ to 70° they are consistently above it. This behaviour completely matches the systematic departures of the hydrogen layer from the principal plane (Fig. 1(*c*)). For the longitude range $l^{\text{I}}=80^\circ$ to 220° we are forced to rely mainly on the less accurate wide-beam surveys (Fig. 1(*a*)). Conclusions drawn from wide-beam surveys alone may be somewhat suspect, but we should at least be sure that the results are consistent with, or can be explained in terms of, other more accurate evidence. Comparing Figs. 1(*a*) and 1(*c*), it is seen that the average position of the wide-beam ridge-lines over most of the 360 degrees of galactic longitude satisfies our test surprisingly well.

The plausibility of the present interpretation of the continuum results is further strengthened by the fact that it is now possible to account for some previously unexplained results. Seeger and Williamson (1951) and Piddington and Trent (1956), on the assumption that the continuum sources were distributed about a plane of symmetry, inferred values for the distance of the Sun from the "galactic plane" much larger than the optical value or that derived from H I observations. Seeger and Williamson's very large values (hundreds of parsecs) have never been properly explained. It has probably been tacitly assumed that wide-beam surveys are suspect anyway, but according to the present interpretation it is the assumption of a plane distribution for the sources which is at fault. The observations appear to be satisfactory, and can be accounted for, at least qualitatively, in terms of distortions in the galactic layer without the need for a large value of the Sun's distance from the plane.

Further support for our hypothesis is given by the observation by Mills (1959) that continuum observations at 3·5 m gave evidence for distinct spiral arms which tended to coincide with H I arms.

The data now available are therefore consistent with the hypothesis that the large-scale spatial distribution of the continuum sources is coincident with that of the neutral hydrogen, the surfaces of maximum density of both being flat in the inner regions of the Galaxy, and systematically distorted in the same way in the outer parts. Further, accepting this hypothesis, the continuum observations in the general direction of the galactic centre can be used to provide independent

evidence on the position of the galactic pole. The parts of the two surveys with the narrowest beams, those of Westerhout and of Hill, Slee and Mills, in the range of longitudes within 60° of the galactic centre were used to obtain new values of Δ, the deviation of the pole from the old one. Because of the limited range of longitudes these observations are ill-suited to a determination of l_0^{I}, the longitude of the pole. Using the method of least squares, the great circle of best fit having a pole at longitude $l_0^{\mathrm{I}} = 347°$ (the H I value) was determined for each of the above surveys. The results are given in Table II. The locations of the pole which correspond to these values of Δ are shown in Fig. 1 of Paper I (p. 129).

TABLE II

Deviations Δ of continuum poles from the old pole, assuming $l_0^{\mathrm{I}} = 347°$

Observer	Frequency (Mc/s)	Longitude range(lI)	$\Delta = 90° - b$I p.e.	Remarks
Westerhout	1390	322° − 27°	1°·55 ± 0°·04	0°·05 greater than H I value
Hill *et al.*	85·5	268° − 10°	1°·42 ± 0°·05	0°·08 less than H I value

It should be borne in mind that fitting a great circle to the observations presupposes that the Sun is in the plane determined by the sources of continuous radiation. From the 21-cm line observations it was found that the Sun is at a height $z_0 = 4\,\mathrm{pc} \pm 12$ above the H I principal plane. If we assume that the average distance of the sources in the region considered is 8 kpc, this value of z_0 would decrease the values of Δ found by 0°·03 ± 0°·09. The quoted probable errors in Δ, given in Table II, are internal errors. A more realistic value would probably be ± 0°·10.

3. *Evidence on the position of the galactic centre*

The use of the direction to the galactic centre for longitude zero has been proposed in the past (see Ohlsson 1932) but was not adopted, probably because of the large uncertainty of this direction at the time.

In the first part of this section we discuss evidence on the direction of the centre derived from broad features of the distribution in longitude of the radio emission. In the second part the evidence on the location of the radio source Sagittarius A is presented. On the basis of the discussion in Paper V we accept this source as being at the centre of the Galaxy.

3.1. *General distribution in longitude.*—For a study of the general distribution of intensity with longitude, only southern hemisphere surveys give an adequate range of longitude on both sides of the centre. For the southern hemisphere continuum surveys from Table I, the intensity along the ridge-line was plotted as a function of galactic longitude. From these curves for the region $l^{\mathrm{I}} = 280°$ through the galactic centre to $l^{\mathrm{I}} = 20°$, *always omitting sections closer to the source Sagittarius A than an aerial beamwidth or so,* the longitudes about which the central maximum was most nearly symmetrical were estimated. These longitudes, together with subjectively estimated probable errors, are listed in Table III. The estimated accuracy of the various determinations is practically independent of the beamwidth. The narrow-beam surveys show more detail but do not give a more accurate determination of the direction of the centre when the smoothed

overall distribution is considered. There are small differences, up to a few degrees, between the values in Table III and those given by the investigators themselves, who have usually given the position of the maximum observed intensity. These latter positions do not agree so well with one another as those obtained by taking into account the distribution in a wide range of longitudes on either side of the centre.

The H I observations may also be used in several ways to determine the direction of the galactic centre from the symmetry of the 21-cm distribution. Kerr (private communication) has made two determinations from a study of the integrated brightness of the 21-cm line profiles for directions in the H I principal plane. In the first method, a centre of symmetry was found from the longitude distribution of the integrated brightness, in a way which was directly analogous to that described above for continuum observations. In the second, the velocity pattern was taken into account by considering the positive and negative velocity parts of the profiles separately. If the Galaxy is symmetrical

TABLE III

Longitude of the galactic centre from general radio distributions

(*a*) *Continuum*

Maximum of intensity distribution in longitude (influence of Sagittarius A removed)

	Frequency (Mc/s)	Beamwidth	l_0 I	est. p.e.
Bolton and Westfold (1950)	100	17°	328°	±3°
Allen and Gum (1950)	100	25°	328°·5	±3°
Piddington and Trent (1956)	600	3°·3	328°	±4°
Hill, Slee and Mills (1958)	85·5	0°·8	327°·5	±2°
			Mean 328°	±1°·5

(*b*) *H-line*

Centre of large-scale H I distribution; Kerr (unpublished)	327°·5	±1°
Rotational centre, large-scale; Kerr (unpublished)	326°	±1°
Rotational centre, small central disk; Oort and Rougoor (Paper V)	327°·70	±0°·10

on a large scale in both structure and velocity, the distribution with longitude of the integrated brightness for positive velocities will be the mirror image about the galactic centre of that for negative velocities.

Consideration of these two forms of symmetry over the central 50° range of longitudes led to the values for the apparent longitude of the galactic centre, tabulated in Table III.

Oort and Rougoor (Paper V) have obtained another value for the centre of rotation of the Galaxy by restricting attention to the small rapidly-rotating disk of neutral hydrogen which they have found in the region of the galactic centre. This value is also tabulated in Table III.

The longitude of the centre, derived from the best optical observations (see Paper V) is in complete agreement with that derived from the radio data.

3.2. *The position of the source Sagittarius A.*—The position of the source Sagittarius A can be determined with a precision far exceeding that with which the galactic centre can be located by other means. Some of the more accurate of the older determinations are shown in Table IV.

Table V gives four new determinations made with high-resolution radio telescopes, each a paraboloid of diameter about 25 metres. These results were generously communicated to us prior to publication. The positions are shown in Fig. 2, with their probable error circles.

An interesting feature of the observations with the highest resolution (Sloanaker, Nichols and McClain, private communication) is that the source shows asymmetry, the position of peak reading being displaced from that of the centre of the half-power positions.

The mean position of the later determinations is $l^{\mathrm{I}} = 327^{\circ}{\cdot}68$, $b^{\mathrm{I}} = -1^{\circ}{\cdot}45$. This position lies within $0^{\circ}{\cdot}01$ in latitude of the "central point" of the adopted H I principal plane, taking into account the displacement of the Sun from the plane. This result, incidentally, is a confirmation of the accuracy of location of the H I principal plane.

It is seen that the longitude of Sagittarius A coincides with that of the galactic centre as determined from general distributions well within the uncertainty of the latter. The latitude of Sagittarius A is also in agreement with that of the appropriate part of the H I principal plane and the continuum ridge-lines.

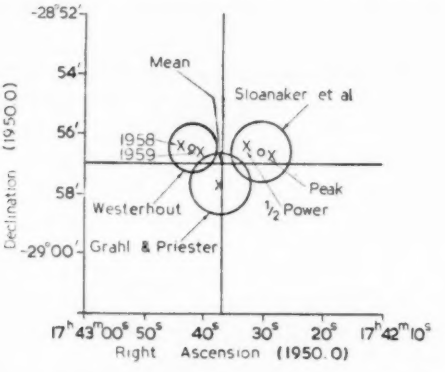


FIG. 2.—*Recent precise determinations of the position of Sagittarius A, with probable error circles (see Table V).*

TABLE IV

Positions of Sagittarius A (older determinations)

Observer	Frequency Mc/s	Beam-width	(1950.0) α	δ
Haddock, Mayer and Sloanaker (1954)	3200	$0^{\circ}{\cdot}42$	$17^{\mathrm{h}}42^{\mathrm{m}}{\cdot}5 \pm 0^{\mathrm{m}}{\cdot}2$	$-29^{\circ}01' \pm 5'$
Hagen, Lilley and McClain (1955)	1400	$0^{\circ}{\cdot}9$	$17^{\mathrm{h}}43^{\mathrm{m}}$	$-28^{\circ}45'$
Estimated from Hill, Slee and Mills (1958)	85·5	$0^{\circ}{\cdot}8$	$17^{\mathrm{h}}43^{\mathrm{m}}{\cdot}3 \pm 0^{\mathrm{m}}{\cdot}7$	$-28^{\circ}48' \pm 12'$
Kerr and Hindman (unpublished)	1400	$1^{\circ}{\cdot}4$	$17^{\mathrm{h}}42^{\mathrm{m}}{\cdot}7 \pm 0^{\mathrm{m}}{\cdot}1$	$-29^{\circ}01 \pm 3'$
Westerhout (unpublished)	1390	$2^{\circ}{\cdot}3$	$17^{\mathrm{h}}42^{\mathrm{m}}{\cdot}1 \pm 0^{\mathrm{m}}{\cdot}8$	$-29^{\circ}00' \pm 15'$
Mean	$\alpha = 17^{\mathrm{h}}42^{\mathrm{m}}{\cdot}7$		$\delta = -28^{\circ}55'$	
	$l = 327^{\circ}{\cdot}72$		$b^{\mathrm{I}} = -1^{\circ}{\cdot}45$	

TABLE V

Positions of Sagittarius A (recent precise determinations made with 25 metre paraboloids)

Observer	Beam-width	Observed position and equinox		Position, reduced to equinox 1950.0		Adopted position	
		α p.e.	δ p.e.	α	δ	α	δ
Westerhout (1958); 1390 Mc/s	0°·57	$17^h43^m10^s.8 \pm 7^s$ (1956.8)	$-28°56'.6 \pm 2'$	$17^h42^m44^s.1$	$-28°56'.4$	$17^h42^m42^s.0$	$-28°56'.5$*
Westerhout (private communication, 1959); 1390 Mc/s	0°·57	$17^h43^m15^s.1 \pm 3^s$ (1959.0)	$-28°56'.9 \pm 0'.7$	$17^h42^m40^s.9$	$-28°56'.6$		
Grahl and Priester (private communication); 1419 Mc/s.	0°·61	$17^h43^m7^s.9 \pm 4^s$ (1958.0)	$-28°57'.9 \pm 1'$	$17^h42^m37^s.4$	$-28°57'.7$	$17^h42^m37^s.4$	$-28°57'.7$
Sloanaker, Nichols and McClain (private communication); 2930 Mc/s.	0°·31 'Peak'	$17^h42^m59^s.5 \pm 5^s$ (1958.0)	$-28°56'.9 \pm 1'$	$17^h42^m29^s.0$	$-28°56'.7$	$17^h42^m30^s.3$	$-28°56'.60$†
	'½ Power'	$17^h43^m03^s.5 \pm 5^s$ (1958.0)	$-28°56'.4 \pm 1'$	$17^h42^m33^s.0$	$-28°56'.4$		

Mean position: $\alpha = 17^h42^m37^s$, $\delta = -28°57'$ (1950.0).

In old galactic coordinates: $l^I = 327°.68$, $b^I = -1°.45$.

*Mean of 1958 and 1959 determinations with weights 1 and 2 respectively.

†Mean of "Peak" and "½ Power" positions with weights 2 and 1 respectively.

Important new information, elucidating the asymmetry found by Sloanaker *et al.* was recently communicated to us by F. D. Drake (private communication). He took exceptionally high resolution observations (25-metre paraboloid at 3 cm, beamwidth 0°·12) and found that the source has a complex structure with several distinct maxima. Since the source is not regular it is not clear which point should coincide with the exact centre of the Galaxy. Hence the use of Sagittarius A as a marker for the centre involves an uncertainty of a few minutes of arc even though its centre of gravity (or peak) may be found with higher precision.

4. *Physical significance of the results*

The outstanding feature of all the observations studied in this paper is the flatness of the disk in the inner parts of the Galaxy. Since it is flat it is probably in an equilibrium position defined by the surface of zero z-component of gravity. The departure from flatness in the external parts is not yet properly explained but the departures exist in regions involving only a small fraction of the total mass of the Galaxy and suggest the influence of forces external to our Galaxy (Kerr 1957; Burke 1957; Kahn and Woltjer 1959). The principal plane defined by the part of the neutral hydrogen not showing these deviations is probably the best reference plane for dynamical studies of the whole Galaxy; the H I principal plane is probably normal to the axis of rotation of the Galaxy.

Apart from its application in defining the galactic pole, the high degree of flatness is of interest in dynamical and evolutionary studies. If the H I disk is an indicator of the surface of zero z-component of gravity, then the latter surface is surprisingly flat. We assume that the Galaxy originated as a condensation in a vast mass of gas, irregularly distributed and having irregular motions with a net angular momentum in the direction of the present axis of rotation, the z-axis. The assumption of irregularity implies that at that stage, the surface of zero z-component of gravity should also have been irregular. As time proceeded, interaction between the various bodies of matter would tend towards an equilibrium state in which, presumably, this surface should be flat and normal to the axis of rotation. For masses of gas the relaxation process should, relative to the age of the Galaxy, be relatively fast owing to frequent collisions; for stars it should be extremely slow, probably too slow for it to have led to equilibrium in the life time of the Galaxy. Since most of the mass in our Galaxy is now in the form of stars, the present-day surface should be an indicator of the general distribution of stars, not of interstellar gas. The interesting point is that despite the slow relaxation time of stars, this surface is now flat. On these assumptions it appears that the relaxation process giving the present mass distribution in the z-direction took place prior to the main phase of star formation in the Galaxy.

The investigation of this aspect of galactic evolution requires both theoretical and observational study. The present evidence, as pointed out in Paper II, is not yet sufficiently precise to define the magnitude of the residual irregularities in the central parts of the neutral hydrogen disk. More refined observations are required before this interesting evolutionary aspect can be properly explored.

The large-scale coincidence in position between the neutral hydrogen and the continuum sources in the galactic disk is an important item of evidence concerning the continuum sources. At the shorter wavelengths, e.g. 20 cm, these are largely H II regions and the association is a natural one. At the longer wavelengths the non-thermal component is predominant. This is currently

believed to be due to synchrotron-type emission from relativistic electrons spiralling in interstellar magnetic fields. Our suggestion is that the determining factor in the coincidence between H I and both types of continuum sources is the concentration of interstellar gas. Observations of 21-cm line emission indicate directly the existence of this concentration; we suggest that an increase of gas density also leads to an increase of continuum emission because it predisposes towards star formation and the release of energy in various ways ranging from the birth of intensely luminous O and B stars, which ionize the gas, to the production of relativistic electrons and magnetic fields.

Accepting the tendency of the continuum and 21-cm sources to be similarly distributed, the question arises whether the various optical objects which show a high galactic concentration show the same departure from a plane as do the H I clouds in the outer parts of the Galaxy. If there were differences in the distributions such evidence might be used to infer long-term changes in the form of the disk. However, there is not yet adequate observational evidence on this point.

5. *Significance of the results in relation to a galactic coordinate system*

While the H I principal plane has an intrinsic interest, it might be considered to be undesirable to use the hydrogen clouds only for a determination of an equator for a new galactic coordinate system. The settling of this question requires a decision on what the coordinate system is meant to achieve. If it is to provide a useful and convenient reference plane for a study of the whole Galaxy and its overall dynamical problems, a plane such as the H I principal plane is desirable. If, on the other hand, it is to be rather a convenient observational datum, then a mean plane fitting a greater variety of the more readily observable objects, say the more accessible optical objects and the radio continuum ridge-lines, may be better. As it turns out, it appears that both these requirements may be satisfied simultaneously. The evidence we have examined together with that discussed in Paper IV indicates that the H I principal plane and its uncertainty fall well inside the range of uncertainty of all other determinations. The use of the H I principal plane as a basis for a coordinate system therefore satisfies all the requirements.

Let us now consider the question of the uncertainties in the position of the galactic pole and centre, and the extent to which these are due to inherent dispersion among the objects under observation (cosmic scatter), or to uncertainties of observation which are likely to be reduced in the near future.

In optical surveys like that of van Tulder, the main observational uncertainty is in the distances. Since the Sun is almost in the plane, errors in distance will have only a second order effect on the position of the pole, although they will show up in the derived height $z_\odot$ of the Sun above the plane. Obscuring clouds will produce apparent irregularities in the distribution of the objects investigated —and will limit the region of the Galaxy which it is possible to survey. Their effect can be regarded as an addition to the cosmic scatter. Optical surveys are most severely limited by obscuration, and not much improvement in present accuracies can be expected.

In the determination of the principal plane of neutral hydrogen, the discussion in Paper II suggested that the observational scatter at present masks any true cosmic scatter which may exist. The reality of the small apparent deviations, of the order of 20 pc, of the hydrogen layer from the principal plane in the flat inner

region cannot be decided at present. Nor should it be forgotten that there are still difficulties in the H-line interpretation—for example, doubts about circular velocity models (see Paper II) which produce uncertainties in the distances. There are also uncertainties in the separation of contributions from "near" and "far" components at the same velocity in the inner region. These uncertainties will presumably be reduced by future optical and radio studies. The improvement hoped for is one in which the observational scatter may be made less than the cosmic scatter. But the task will be a major one and a substantial improvement is likely to take many years.

In the case of the longitude of the galactic centre two different orders of accuracy are involved, that of the direction to the radio source Sagittarius A, which we accept as being at the galactic centre (Paper V), and that of the position of the "centre of gravity" of the general distribution over a range of longitude on either side of the galactic centre. It is fortunate that the better determined position of Sagittarius A falls well within the uncertainty of the direction to the galactic centre derived from general distributions. In this case it is not necessary to raise the question whether such a "nucleus" is an appropriate centre for studies of the Galaxy as a whole. Even in the extreme and most unlikely case that Sagittarius A is not even in the general vicinity of the galactic centre the use of the direction to Sagittarius A will not have degraded our location of the direction to the centre. On the other hand, the discovery that the source is complex in shape prevents us from using the full precision of the radio observations. We require more physical information on the nature of the nucleus of a galaxy before we can decide what part of Sagittarius A should be taken as the galactic centre.

Finally, reverting to the question of what the coordinate system is meant to achieve, it would appear that for the purpose of providing a convenient display of objects concentrated towards the galactic plane, a system is adequately well placed when its uncertainty in position is a small fraction of the spread of objects to be displayed. From what has been discussed it appears that the H I principal plane and the adopted position of the galactic centre well satisfy this requirement for both optical and radio continuum objects.

Acknowledgments.—We gratefully acknowledge the assistance we have received from many astronomers—in particular from G. Westerhout; R. M. Sloanaker, J. H. Nichols and E. F. McClain; B. H. Grahl and W. Priester; F. D. Drake; and F. G. Smith—for providing us with unpublished information—and from F. J. Kerr for assistance throughout our studies.

C.S.I.R.O. Radiophysics Laboratory,
Sydney,
Australia.

1960 *March* 16

References

Allen, C. W. and Gum, C. S., 1950, *Aust. J. Sci. Res.*, **A3**, 224.
Baldwin, J. E., 1955, *M.N.*, **115**, 684.
Blythe, J. H., 1957, *M.N.*, **117**, 652.
Bolton, J. G. and Westfold, K. C., 1950, *Aust. J. Sci. Res.*, **A3**, 19.
Burke, B. F., 1957, *A.J.*, **62**, 90.
Denisse, J. F., Leroux, E. and Steinberg, J. L., 1955, *Comptes Rendus, Acad. Sci., Paris*, **240**, 278.

Dröge, F. and Priester, W., 1956 *Zs. f. Astrophys.*, **40**, 236.
Haddock, F. T., Mayer, C. H. and Sloanaker, R. M., 1954, *Nature*, **174**, 176.
Hagen, J. P., Lilley, A. E. and McClain, E. F., 1955, *Ap. J.*, **122**, 361.
Hanbury Brown, R. and Hazard, C., 1953, *M.N.*, **113**, 109.
Hill, E. R., Slee, O. B., and Mills, B.Y., 1958, *Aust. J. Phys.*, **11**, 530.
Kahn, F. D. and Woltjer, L., 1959, *Ap. J.*, **130**, 705.
Kerr, F. J., Hindman, J. V. and Carpenter, M. S., 1957, *Nature*, **180**, 677.
Kerr, F. J., 1957, *A.J.*, **62**, 93.
Kirillova, T. S., 1955, *A.J.* (U.S.S.R.), **32** (No. 2), 192.
Ko, H. C. and Kraus, J. D., 1957, *Sky and Tel.*, **16**, 160.
McGee, R. X., Slee, O. B. and Stanley, G. J., 1955, *Aust. J. Phys.*, **8**, 347.
Mills, B. Y., 1959, *Symp. I.A.U.*, **9**, 431.
Muller, C. A. and Westerhout, G., 1957, *B.A.N.*, **475**, 151.
Ohlsson, J., 1932, *Ann. Lund. Obs.*, **3**.
Piddington, J. H. and Trent, G. H., 1956, *Aust. J. Phys.*, **9**, 481.
Reber, G., 1948, *Proc. Inst. Radio Engrs. N.Y.*, **36**, 1215.
Schmidt, M., 1957, *B.A.N.*, **475**, 247.
Seeger, C. L. and Williamson, R. E., 1951, *Ap. J.*, **113**, 21.
Westerhout, G., 1957, *B.A.N.*, **475**, 201.
Westerhout, G., 1958, *B.A.N.*, **488**, 215.

OPTICAL DETERMINATIONS OF THE GALACTIC POLE

(PAPER IV)

A. Blaauw

(Received 1960 March 21)

Summary

A discussion is given of two new determinations of the galactic pole, based on recent surveys of distant OB stars and Cepheids (Table I). It is shown that, within the observational uncertainty, the plane defined by all optical objects observed between 1000 and 4000 pc from the Sun is parallel to the plane defined by the neutral hydrogen.

It is concluded, however, that the optical objects should not be used in the determination of a new galactic pole because (i) they represent a very limited sample of the Galaxy, (ii) they are not free from the influence of observational selection (absorption), (iii) they are partly situated in the outer distorted regions of the Galaxy and (iv) they are closely related to the neutral hydrogen and therefore not really independent.

---

1. *Introduction*

From a study of the distribution of objects of strong galactic concentration and planetary nebulae, and from the directions of the motions of high velocity stars, van Tulder (1942) found the galactic pole at

$$\begin{aligned} \alpha &= 12^{h}\,44^{m}\cdot0 \pm 1^{m}\cdot1 \text{ (m.e.)} \\ \delta &= +27\cdot5^{\circ} \pm 0^{\circ}\cdot2 \text{ (m.e.)} \end{aligned} \qquad (1900)$$

or

$$l^{I} = 330^{\circ} \quad b^{I} = +89^{\circ}\cdot0\dagger,$$

deviating about $1^{\circ}\cdot0$ from that adopted by Ohlsson for the Lund Tables. His study embraced, among others, supergiant stars, O and B-type stars, Cepheids and open clusters, up to distances of 4000 pc. The great mass of stars of the Milky Way did not enter into his solutions. Several earlier determinations of the galactic pole, to which references may be found in van Tulder's article, were based on the apparent distribution of the stars in general. Justification of van Tulder's approach lies in the recognition that the irregularly distributed, nearby obscuring clouds greatly influence the apparent distribution of the stars.

More recent determinations of the pole than van Tulder's have been made along the same lines, using Cepheid variables, for instance by Ashbrook and Duncombe (1952) and by Kirillova (1955). The latter author introduced the assumption that the galactic plane, apart from fitting the Cepheids, should also pass through the adopted position of the galactic centre and through the Sun, and she incorporated evidence from radio observations.

In order to broaden somewhat the basis of our report, we have made additional solutions for the galactic pole using the recent photometric observations of Cepheids published by Walraven, Muller and Oosterhoff (1958) and Hiltner's

†The galactic coordinates used in this paper are the old system.

(1956) photometry of O and B-type stars. The following section briefly describes these solutions; they are discussed with the other solutions in the third section.

2. *New solutions from O and B-type stars and Cepheids*

2.1. *O and B-type stars.*—Distances were derived from the distance moduli as given by Hiltner, which are corrected for interstellar absorption. The stars of Hiltner's list had been selected from objective prism surveys at Cleveland and Tonantzintla. These surveys were confined to the zone between $\pm 6°$ latitude and accordingly Hiltner's list does not contain stars beyond these limits. Our galactic pole solutions were confined to the stars at distances larger than 2500 pc. At this distance the 6° latitude limit corresponds to a distance from the galactic plane, $z = 260$ pc; for smaller distances from the Sun this limiting z is proportionately smaller and forms a cut-off which might affect the solution. Hiltner's list does not contain all the stars listed in the Cleveland and Tonantzintla surveys and it is not clear to what extent his choice may have influenced the galactic pole solution; we have assumed this to be negligible. A small number of stars which are at large distances from the galactic plane, as found from a provisional solution, were subsequently excluded. Table I gives the results found from two solutions based on the 291 remaining stars.

TABLE I

Longitudes, l_0^{I} and latitudes, $b_0^{\mathrm{I}} = 90° - \Delta$ of the galactic poles and distances z_0 of the Sun from the galactic plane, as derived from OB-type stars and Cepheids

	OB-type stars beyond 2000 pc		
	Solution A All longitude groups equal weight	Solution B Weights proportional to numbers of stars in most longitude groups	Cepheids
l_0^{I}	335° ±12° (p.e.)	353° ±10° (p.e.)	326° ±8° (p.e.)
$\Delta = 90° - b_0^{\mathrm{I}}$	1°·9 ± 0°.35	1°·2 ± 0°·24	1°·31 ± 0°.18
z_0	+18 pc ± 18	+2 pc ± 9	+27 pc ± 5

In Solution A equal weights were assigned to longitude intervals of 10 degrees in order to eliminate entirely the effect of the accumulation of the stars in some longitudes, particularly those between $l^{\mathrm{I}} = 70°$ and 110°. This solution is given in the second column of Table I.

In Solution B weights were given proportional to the numbers of stars with the exception of the interval of longitudes 70° to 110°, where the maximum value of the weight occurring in the other intervals was adopted. This solution is given in the third column of Table I.

2.2. *Cepheids.*—From the list of Cepheids (mostly southern) published by Walraven, Muller and Oosterhoff, we excluded Population II Cepheids and all other objects at distances z from the galactic plane exceeding 200 pc. Of the remaining 157 stars, 27 per cent are within one kiloparsec from the Sun, 37 per

cent between one and two kiloparsec, 22 per cent between two and three kiloparsec and 15 per cent beyond three kiloparsec. The result of the solution, in which all stars were given equal weight, is in the fourth column of Table I.

The poles found from these three solutions are quite close to that of the H I principal plane as given in Paper II. The differences cannot be considered significant in view of their probable errors. These results are incorporated in the compilation of Section 3.

The general distribution of the OB stars and Cepheids with respect to the new galactic plane is illustrated in Figures 1 and 2. We first computed distances z with respect to this plane (which passes through the Sun) and next mean values of z for areas in the plane limited by the distances 1000 pc, 1500 pc, 2000 pc, etc. and the longitudes 0°, 30°, 60°, ... These mean values are marked in the figures. Mean values of z derived from less than ten stars are in parentheses. For the OB stars (Fig. 1) we do not present the figures within 1000 pc since these might be biased because of the cut-off mentioned before. For the region between 1000 and 2000 pc distance, this will be the case to a smaller degree. For the Cepheids averages were formed for a combination of areas in order to avoid figures based on less than 3 stars.

The preponderance of negative values in Fig. 2 (Cepheids) is in agreement with the large value for the distance of the Sun above the plane for these objects, which is given in the last line of Table I. Fig. 1 (OB stars) shows much larger fluctuations between consecutive longitude intervals. These are partly due to the presence of local groupings like the h and χ Persei association around 100° longitude, which extends to $-5°$ latitude. The positive values in the longitude interval 30° to 60° cannot be so simply ascribed to one large concentration. The fact that there is so much resemblance in the sign in various distance intervals for each of the longitude intervals 30° to 60°, 60° to 90° and 90° to 120° is probably due to uncertainties in the distance estimates of the individual stars, which tend to spread stars which in reality are close together (clustering) along the radial direction. Interstellar absorption may also play an important role in this respect.

On the whole, the distribution of the Cepheids seems to be more even than that of the OB stars, a phenomenon which would merit further study. Such a study should then also comprise various other kinds of objects. This is, however, beyond the scope of the present contribution.

3. *Assessment of the data*

van Tulder's paper contains a diagram (his Fig. 1) representing the positions of the poles for the various groups analysed by him, which exhibits the general agreement of the well-determined poles with the finally-adopted average pole.

In Fig. 3 we have given a somewhat different representation of van Tulder's results. The figure is adapted from Gum and Pawsey (1958). It illustrates the progressive decrease of the scatter of the poles for groups of objects at increasing distance from the Sun, and their concentration towards the pole of the H I principal plane. Poles for objects at distances unknown to van Tulder have been omitted. The distance groups $\leqslant 500$ pc, 500 to 1000 pc and 1000 to 2000 pc contain van Tulder's solutions from O and B-type stars, supergiants, Cepheids and open clusters. His distance groups at 2700 pc and 3300 pc refer to Cepheids and open clusters, respectively. We have also plotted Ashbrook and Duncombe's Cepheid solution IV (Cepheids brighter than photographic

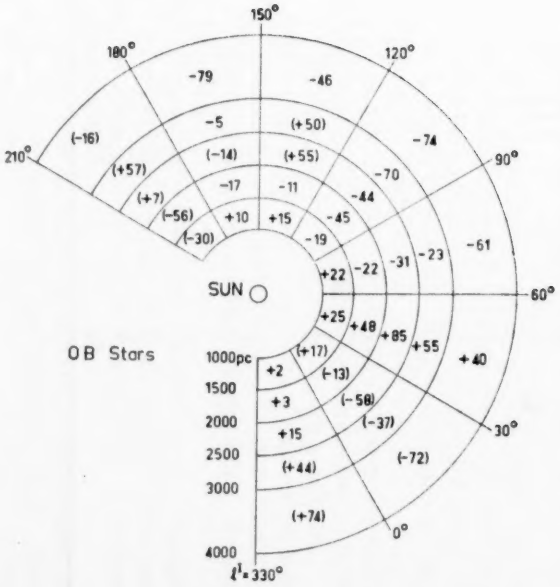


FIG. 1

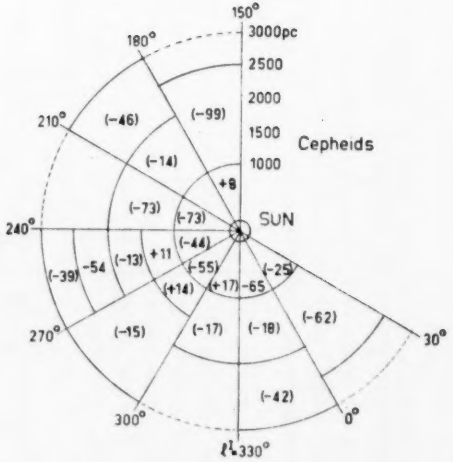


FIG. 2

Mean distances from the new galactic plane (which passes through the Sun), expressed in parsecs, for the OB stars (Fig. 1) and the Cepheids (Fig. 2), treated in our new solutions. Values in parentheses are based on fewer than 10 stars.

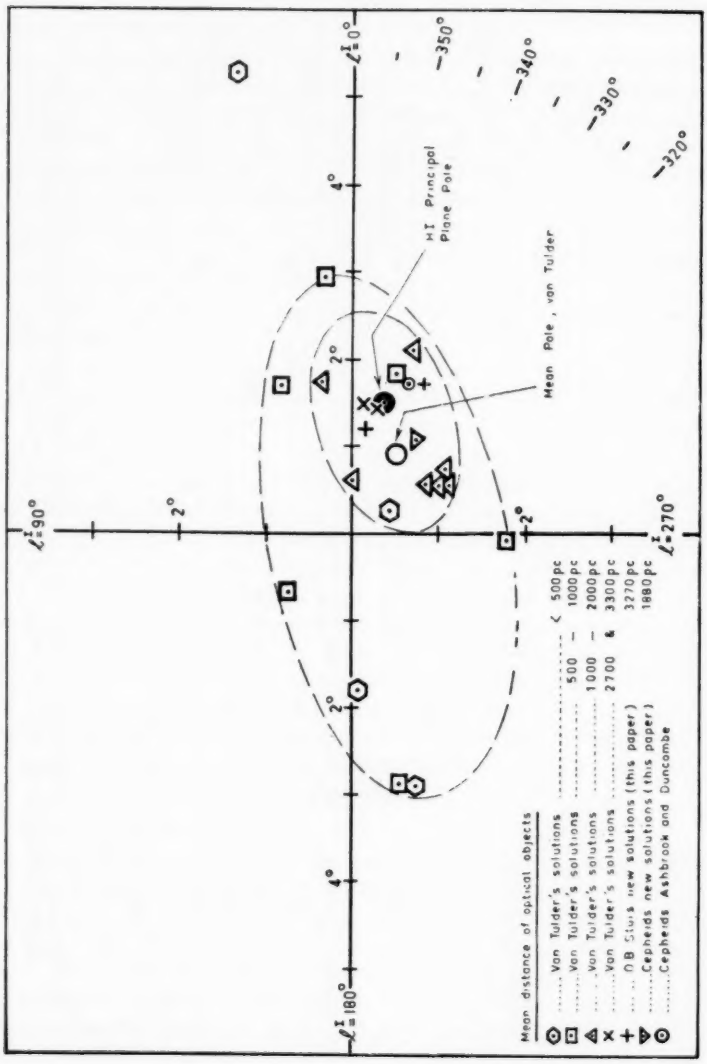


FIG. 3.—*Poles derived from optical objects at different distances from the Sun. The oval curves include the poles derived from objects up to* 1000 *and* 2000 *pc, respectively.*

magnitude 12·5) which according to these authors is the best compromise between increasing numbers of stars with decreasing brightness and increasing effects of observational selection. We have added the results of our solutions for Cepheids and OB stars.

Apart from the decreasing scatter, a progressive movement of the mean pole is apparent in Fig. 3 if one considers the centre of gravity of the areas enclosed by the oval curves. These curves are drawn to include the poles from objects up to 1000 and 2000 pc, respectively. The poles of our Cepheid and OB star solutions appear to lie very close to those previously determined for such objects and to the H I principal plane pole.

Considering now all available data, we arrive at the following conclusions:

(i) Within the whole region accessible to the optical observations (i.e. up to 4000 pc) the large-scale distribution of the objects studied defines a plane, parallel to the H I principal plane within the uncertainty of the determination. The pole of this plane, as determined from all objects in van Tulder's list beyond 1000 pc, and including Ashbrook and Duncombe's solution IV and the determinations from the preceding section, is found to be

$$l_0^{\mathrm{I}} = 334^\circ \pm 5^\circ \text{ (p.e.)}, \quad \Delta = 90^\circ - b_0^{\mathrm{I}} = 1^\circ{\cdot}41 \pm 0^\circ{\cdot}26 \text{ (p.e.)}.$$

The distance of the Sun from this plane is $+22\text{pc} \pm 2$ (p.e.). This would indicate an average position of the 'optical' plane somewhat below the H I principal plane, but this shift can hardly be considered as well established since the probable error given is an internal probable error only. In view of the fact that a certain proportion of the objects used for the determination are situated in the outer distorted parts of the Galaxy (compare Fig. 1 in Paper II, p. 133), such a shift is not unexpected.

(ii) More locally, i.e. within 1000 pc, the average distribution seems to follow a plane somewhat inclined to the H I principal plane, the tilt being of the order of $1^\circ{\cdot}5$, with the pole moving approximately towards the direction $l^{\mathrm{I}} = 180^\circ$. It is of some interest to notice that this tilt is in the same direction as, though much less than, the tilt of the Gould belt.

(iii) The mean pole derived from the optical data beyond 1000 pc from the Sun has an uncertainty of about $0^\circ{\cdot}25$. This is somewhat more than the uncertainty of the pole of the H I principal plane, but not greatly so.

While thus the optical evidence might be considered as being of value in defining the new galactic pole, we have nevertheless decided to discard it entirely in favour of the evidence from the neutral hydrogen data. The reasons for doing so are the following:

(*a*) The H I principal plane represents the neutral hydrogen distribution in a region symmetrically situated around the galactic centre and many times larger than that covered by the optical observations. It is therefore of more fundamental significance.

(*b*) The optical determinations are not free from influence of observational selection of the stars caused by the interstellar absorption. No such selection affects the H I solution.

(*c*) Between a quarter and a half of the region of the Galaxy on which the optical determinations are based is situated in the "distorted" outer parts, which were excluded from the H I solutions.

(*d*) The optical objects all belong to the young Population I and are physically closely related to the neutral hydrogen. They do not, therefore, represent evidence of an entirely independent nature. Furthermore, they represent a total mass which is a small fraction of that represented by the neutral hydrogen.

For the present we do not possess very accurate optical data with reference to the galactic plane for any entirely different part of the galactic population. Oort (1942) has determined the pole from counts of stars at large distances (up to 1500 pc) from the galactic plane and found $l^{\text{II}} = 260^\circ{\cdot}0$, $\Delta = 0^\circ{\cdot}6$ with a probable error of $0^\circ{\cdot}7$. This differs from the H I principal plane pole by about $1^\circ{\cdot}6$. It would be most valuable to check the pole by means of a component of the disk Population II at still larger distances. In principle, this might be done by means of counts on objective prism plates of faint red stars in the near infrared, which reach the periphery of the galactic disk.

It is an interesting, and in some respects convenient circumstance, that the local "optical" plane coincides so well with the H I principal plane—a circumstance which might not have been anticipated in view of the large deviations of the neutral hydrogen from this plane in regions of the Galaxy at distances from the centre comparable with that of the Sun, but at other galactocentric longitudes.

Kapteyn Astronomical Laboratory,
Groningen,
The Netherlands.

1960 *March* 16

References

Ashbrook, J., and Duncombe, R. L., 1952, *A.J.*, **56**, 204.
Gum, C. S., and Pawsey, J. L., 1958, C.S.I.R.O. Radiophysics Lab. Rept. 137.
Hiltner, W. A., 1956, *Ap. J. Suppl.*, **2**, 389.
Kirillova, T. S., 1955, *Astr. J. U.S.S.R.*, **32**, 192.
Oort, J. H., 1942, *B.A.N.*, **9**, 324.
van Tulder, J. J. M., 1942, *B.A.N.*, **9**, 315.
Walraven, Th., Muller, A. B., and Oosterhoff, P. Th., 1958, *B.A.N.*, **14**, 81.

THE POSITION OF THE GALACTIC CENTRE

(PAPER V)

J. H. Oort and G. W. Rougoor

(Received 1960 March 21)

Summary

The direction to the radio source Sagittarius A is found to be within 0°·03 of the direction to the galactic centre as determined from a number of precise optical and radio observations, including the recently discovered rapidly rotating disk of neutral hydrogen around the centre. This, together with the fact that the source is unique among the known sources, makes it highly probable that Sgr A is situated at the centre. Additional evidence comes from the observation of 21-cm absorption lines, which indicate that the distance of Sgr A is equal to or greater than the distance to the galactic centre.

Objects of Baade's Population II show a very strong concentration towards the centre of the Galaxy. Observations of RR Lyrae variables in a field centred at $l^{\text{I}} = 328°{\cdot}7$, $b^{\text{I}} = -5°{\cdot}3$* show that at a distance of 600 pc from the centre the density of these variables is about 700 times that in our vicinity (Baade 1958). Similar concentrations are observed for the planetary nebulae, which form a typical old disk population. The density near the nucleus must in this case be at least 1000 times that near the Sun.

The actual centre is hidden by absorbing clouds which must have a very large optical depth. If we were able to penetrate through this matter, we would probably see a well-defined small nucleus, such as is shown by spirals of comparable type. In the Andromeda nebula, for example, a dense nucleus is observed, with a diameter of about 2″·5, or 7 pc (Baade 1955). So far, no optical observations have succeeded in distinguishing a sharp nucleus in the Galaxy. It would seem very important to search for this at a wavelength around 2 microns, where the penetrating power should be sufficient.

For finding the direction of the centre from optical observations we must, for the present, rely on regions more than 2° away from the centre, where the absorption has become sufficiently small to show planetary nebulae. The best value for the *longitude* at present available is that derived from surveys of planetary nebulae. Minkowski (unpublished; for a preliminary discussion cf. Minkowski 1951) finds $l^{\text{I}} = 327°{\cdot}7$. The uncertainty may be of the order of 0°·5; it is mainly caused by the unevenness of the absorption. The result agrees well with the less accurate value obtained previously from globular clusters, viz. $l^{\text{I}} = 327° \pm 2°$ (estimated probable error) (Shapley 1930).

Because of the absorption the *latitude* cannot be determined with much accuracy from these objects. It can, however, be derived with considerable precision from the situation of the layer of neutral hydrogen gas as found from 21-cm line measurements. From the data in Paper II, the centre of the H I principal plane is found to be at $b^{\text{I}} = -1°{\cdot}44$. It can also be determined from

* The galactic coordinates used in this paper are the old system.

the continuous radiation in the region within 50° or 60° from the longitude of the centre. The mean value of the latitude of the continuum "ridge lines" in the centre direction, derived from the data in Table II of Paper III, is $b^{I} = -1°{\cdot}40$.

A still greater precision might possibly be obtained by using only the hydrogen in the region within 3 kpc from the centre, which can, at least in part, be recognized by its high velocity; but the discussion of this has not yet been finished.

The authors have recently adduced evidence (Rougoor and Oort 1960) of a concentration of neutral hydrogen quite close to the centre (within about 500 pc), which forms a thin disk rotating at a high velocity. The discussion of the extensive observational material obtained at Dwingeloo is not yet entirely completed. A provisional reduction has given a mean latitude of $b^{I} = -1°{\cdot}48 \pm 0°{\cdot}03$ (estimated probable error) for the central plane of the small disk, which is likely to contain the mass centre of the Galaxy. The same disk gives also independent accurate information on the *longitude* of the centre of rotation, which is found to be $l^{I} = 327°{\cdot}70 \pm 0°{\cdot}10$ (estimated probable error).

Other dynamical evidence concerning the longitude of the centre, such as may be derived from stars in the neighbourhood of the Sun, or from the motions of globular clusters, is consistent with the above evidence, but is of so much lower accuracy that there is no point in considering it further for the present purpose.

The dynamical evidence from the small rotating disk at the centre agrees very well with that given by the planetary nebulae. Independently from the evidence furnished by the radio source Sagittarius A—which we shall consider below—we may adopt, from the data discussed above, the following position for the mass-centre of the Galaxy: $l^{I} = 327°{\cdot}70 \pm 0°{\cdot}10$, $b^{I} = -1°{\cdot}46 \pm 0°{\cdot}05$, the errors being estimated probable errors.

The position of Sagittarius A has been discussed in Paper III and the values of the coordinates finally adopted are $l^{I} = 327°{\cdot}68$, $b^{I} = -1°{\cdot}45$. This position agrees so precisely with the direction of the galactic centre as found above, that this by itself makes it almost certain that Sgr A is situated *at* the centre of our Galaxy. For Sgr A is not only one of the five brightest sources, but it is also unique among the known sources, consisting, as it does, of a small, apparently thermal core surrounded by a more extensive non-thermal envelope (Westerhout 1958). It would be an extremely improbable coincidence if this unique source should accidentally lie within 0°·1 of the centre without being connected with it. The interesting recent discovery by Drake (1959) that the thermal source is in reality multiple, does not make any difference to this conclusion.

Though the evidence for the connection between Sgr A and the galactic centre provided by the positions in the sky is sufficiently convincing, it remains of some interest to inquire whether there is direct evidence that also the *distance* of Sgr A is of the right order.

That Sgr A is situated at a large distance from the Sun is shown by the 21-cm absorption lines observed in the spectrum of this source. The most striking evidence comes from a spiral arm whose distance from the centre is estimated to be about 3 kpc (van Woerden, Rougoor and Oort 1957; Oort and Rougoor 1958; Rougoor and Oort 1959). This arm, which has been named the "3-kpc expanding arm", passes between us and the centre. It has a remarkable outward motion in addition to partaking in the general rotation of the Galaxy. At the point where it passes Sgr A it shows a strong absorption line at a velocity which coincides with the emission velocities observed in the "3-kpc expanding arm" on either

side of the longitude of the centre. The occurrence of this absorption shows that Sgr A lies behind the arm and that it is, therefore, further than 5 or 6 kpc from the Sun. There are also absorption lines at higher negative velocities which may be due to the small rotating disk at the centre. If this interpretation is correct it would indicate that the distance of the radio source is equal to that of the centre, or greater.

In view of the evidence given it seems fairly safe to assume that Sagittarius A can be identified with the galactic centre.

Sterrewacht,
Leiden,
The Netherlands:
1960 March 16.

References

Baade, W., 1955, *Mitt. Astr. Ges.*, 51–59.
Baade, W., 1958, *Stellar Populations*, edited by D. J. K. O'Connell, S.J., Specola Astr. Vaticana, Ric. Astr., **5**, 310.
Drake, F. D., 1959, *A.J.*, **64**, 329.
Minkowski, R., 1951, *Publ. Obs. Univ. of Michigan*, **10**, 25.
Oort, J. H., and Rougoor, G. W., 1958, *Acad. Sci. Amsterdam Versl.*, **67**, 139.
Rougoor, G. W., and Oort, J. H., 1959, *I.A.U. Symposium No.* 9, 416 (Paris Symp. on Radio Astronomy).
Rougoor, G. W., and Oort, J. H., 1960, *Proc. Nat. Ac. Sci. Washington*, **46**, No. 1.
Shapley, H., 1930, *Star Clusters*, McGraw-Hill, New York, 22.
van Woerden, H., Rougoor, G. W., and Oort, J. H., 1957, *C.R.*, **244**, 1691.
Westerhout, G., 1958, *B.A.N.*, **14**, 215 (No. 488).

THE VARIATION OF METEOR HEIGHTS WITH VELOCITY AND MAGNITUDE

J. S. Greenhow and J. E. Hall

(Communicated by the Director, Jodrell Bank Experimental Station, University of Manchester)

(Received 1960 March 22)

Summary

The variation of the heights of occurrence of meteors, with velocity and magnitude, are compared with the theoretical predictions. The photographic observations are shown to be in good agreement with theory over the range of zenithal visual magnitudes from -3 to $+3$ and for velocities of 13 to 70 km/sec. Very large departures from the predicted heights are found for radio echo observations of $+6$ magnitude meteors. Reasons for these departures are discussed.

---

1. *Introduction*

When meteors enter the Earth's upper atmosphere, evaporated meteor atoms become excited and ionized by collisions with air molecules. The height at which a meteor vaporizes varies with its mass m and velocity V, and theories have been developed which relate the rate of evaporation to m, V, other physical properties of the meteor, and the properties of the upper atmosphere. The heights and velocities of meteors brighter than $+3$ zenithal visual magnitude have been determined photographically, while meteors of about $+6$ magnitude have been investigated by radio echo means. It is the purpose of this paper to compare the experimental measurements with the theoretical predictions of the variation of height with velocity, luminosity and electron density. In particular, attention is drawn to the serious anomalies which occur for the radio echo observations.

2. *Variation of meteor height with velocity, electron density and luminosity*

(*a*) *Theory.*—The atmospheric pressure, $p_{\max}$, at the height at which a meteor undergoes its maximum rate of evaporation, has been determined as a function of the initial mass of the meteor m_∞, the meteor velocity V, and the density of the meteoric material ρ_m (Whipple 1943, Herlofson 1948). Other factors which depend on the physical properties of the meteor enter into the equation for $p_{\max}$, but these are not likely to vary if a representative group of meteors is examined. $p_{\max}$ is given by the expression

$$p_{\max} \propto \frac{m_\infty^{1/3} \rho_m^{2/3}}{V^2}. \qquad (1)$$

A number of simplifying assumptions are made in deriving this expression, and Weiss (1958) has shown that for very slow meteors the denominator should be multiplied by a correction term of the form $(1 + 10l/V^2)$, where l is the latent heat of vaporization of the meteor material.

As V varies, a photographic system selects meteors which emit a particular amount of light energy L per unit path length; thus, in order to compare

equation (1) with the results of experiments, it is necessary to replace the mass by a luminosity term.

If a fraction τ of the kinetic energy of a meteor appears in the form of light, then the rate of emission of light energy, I, is given by

$$I = -\frac{1}{2}\frac{dm}{dt}\tau V^2 \propto nV^3\tau. \tag{2}$$

n is the number of atoms evaporated per unit length of trail. On the assumption that $\tau = \tau_0 V^a$, where τ_0 is a constant, equation (2) becomes

$$I \propto nV^{3+a} \tag{3}$$

giving

$$L \propto nV^{2+a}. \tag{4}$$

At the point of maximum rate of evaporation $n \propto m_\infty$, and assuming $\rho_m =$ constant, we have

$$p_{\max} \propto \frac{L_{\max}{}^{1/3}}{V^{8/3+a/3}}. \tag{5}$$

Similarly a radio echo equipment observes meteors with a particular value of the linear electron density α per unit length, for all velocity groups, where

$$\alpha \propto \frac{dm}{dt}\frac{\beta}{V}, \tag{6}$$

and β is the probability that an evaporated meteor atom will produce an electron by collision processes.

Assuming $\beta = \beta_0 V^b$, where β_0 is a constant, equation (6) becomes

$$\alpha \propto nV^b \propto m_\infty V^b, \tag{7}$$

and for the variation of $p_{\max}$ with velocity for a radio echo equipment we obtain

$$p_{\max} \propto \frac{\alpha_{\max}{}^{1/3}}{V^{2+b/3}}. \tag{8}$$

(*b*) *Variation of τ and β with velocity.*—The variation of luminous efficiency has been considered theoretically by Öpik (1933), who shows that τ is approximately proportional to the first power of velocity. A value $a = 1$ has been assumed in the photographic meteor work at Harvard (Whipple 1943, Jacchia 1952). In a later theoretical paper Öpik (1958) shows that a varies between -1 and $+0{\cdot}6$ and from photographic work in Czechoslovakia, Ceplecha (1958), finds that τ is very nearly independent of velocity. As a enters equation (5) only as $a/3$, its exact value is not critical. A value $a = 1$ will therefore be assumed.

The variation of β with velocity is more uncertain. Evans (1954) assumes that β is independent of velocity, and shows that the observed distributions of meteor heights are consistent with $b = 0$. However, if $\beta =$ const, and as a meteor of $+6{\cdot}0$ mag. produces about 10^{12} electrons/cm at a velocity of 60 km/sec (Greenhow and Hawkins 1952), it follows that at a velocity of 15 km/sec this value of α would be produced by a very faint meteor of only $+11$ mag. Therefore the velocity distribution of radio meteors should show a much higher proportion of slow meteors than do the visual and photographic distributions. This is not observed (Whipple 1955).

It will be shown in this paper, and in a later paper, that previous radio echo measurements of meteor height–velocity distributions and rates do not give the true heights and rates. It is apparent, however, that the number–velocity

distributions for radio echo observations do not differ in order of magnitude from the photographic observations. Similar rates as a function of velocity would imply that

$$L/a \propto V^{2+a-b} = \text{const} \qquad \text{(eqns. (4), (7))} \tag{9}$$

which, for $a = 1$, gives $\beta \propto V^3$. b is therefore of the order of 3.

In principle the ratio τ/β can be determined from combined visual (or photographic) and radio echo observations. In practice this is not easy because of the difficulty of interpreting the radio echo data in terms of the electron line densities in the meteor trails. The most recent interpretation of the combined visual and radio echo data (Davis, Greenhow and Hall 1959), allowing for the effect of electron attachment on enduring echoes, shows that the ratio of luminous efficiency to ionizing efficiency J is independent of velocity, indicating that $\beta \propto V^3$.

As much of the light in meteor spectra is due to ionized meteor atoms (Cook and Millman 1955), it is reasonable to suppose that light and ionization are roughly proportional. This argument also leads to $\beta \propto V^3$. A value of $b = 3$ will therefore be adopted.

We have, therefore, for the atmospheric pressure at the height at which a meteor emits its maximum amount of light energy per unit length, and produces the maximum amount of ionization per unit length

$$p_{\max} \propto L_{\max}{}^{1/3}/V^3, \tag{10}$$

$$p_{\max} \propto \alpha_{\max}{}^{1/3}/V^3. \tag{11}$$

These theoretical relationships will now be compared with the available experimental observations.

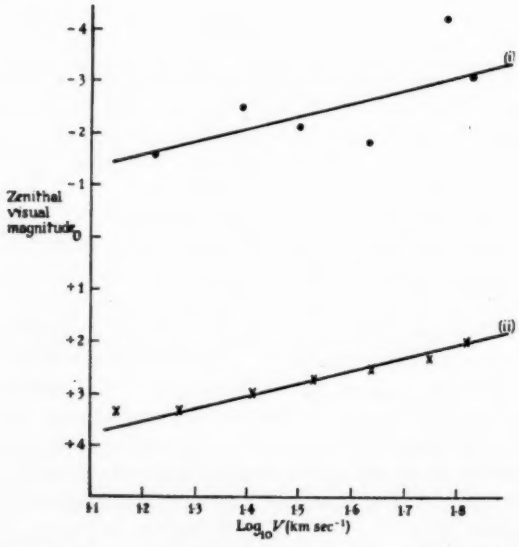


FIG. 1.—*The variation of mean visual magnitude with velocity for faint and bright photographic meteors. The points are based on data published by Jacchia* (1952), *and Hawkins and Southworth* (1958).

3. *Observed height–velocity relationships*

(*a*) *Bright photographic meteors.*—Photographic observations of very bright visual meteors were carried out over a long period at Harvard College Observatory, and a comprehensive list of meteor data has been published by Jacchia (1952). For the purpose of comparison with other observations, the sporadic meteors have been selected from the photographic results, and divided into a number of velocity groups. The variation of zenithal visual magnitude with log velocity for these meteors is shown in Fig. 1 (i). The straight line fitted to the observations has the theoretical slope, assuming that a photographic system observes meteors of constant L in all velocity groups. The mean magnitude varies from $-1{\cdot}4$ for meteors with a velocity of 13 km/sec to $-3{\cdot}2$ for 70 km/sec meteors.

The variation of mean height of maximum light with log velocity for these meteors is shown in Fig. 2 (i). As we are primarily interested in the variation of atmospheric pressure with velocity it is necessary to convert the ordinate in Fig. 2 to a pressure scale. This can be done on the assumption of an isothermal atmosphere, where

$$p \propto \exp(-h/H).$$

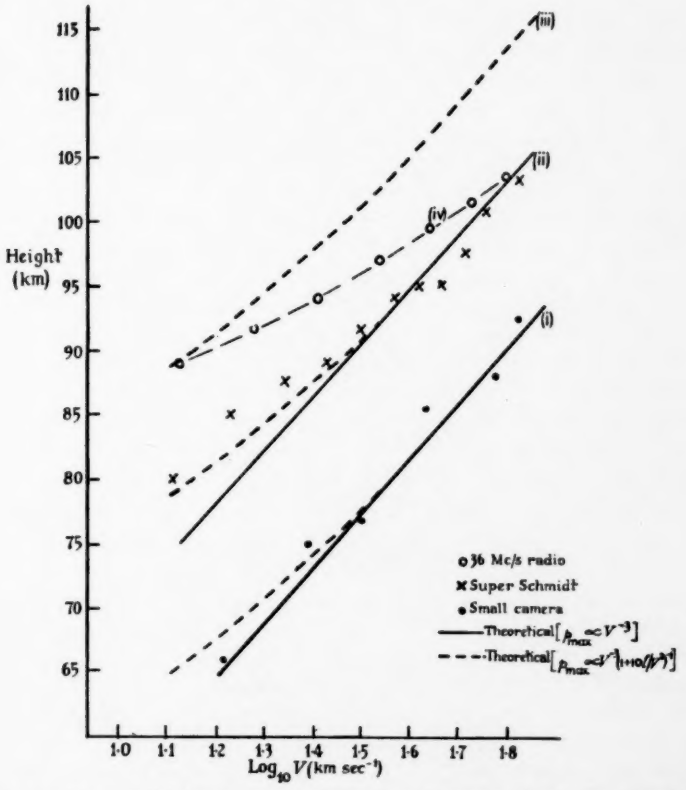


FIG. 2.—*The variations of the heights of meteors with velocity, for three magnitude groups. The photographic and radio echo observations are compared with the theoretical predictions.*

h is the height above some arbitrary zero, and H the atmospheric scale height. A mean value of $H=7$ km has been assumed. This is probably too high at the lowest heights, and too low at the greatest heights. The differences of about 1 km are not important in the present discussion. The ordinate in Fig. 2 is therefore essentially a linear scale in log pressure, the pressure changing by a factor 2·7 every 7 km.

A straight line with a slope corresponding to $p_{\max} \propto V^{-3}$ is shown for comparison with the experimental points in Fig. 2. There is good agreement between theory and experiment. A curved line, including the correction term derived by Weiss, is also shown but the scatter on the experimental points is too great to distinguish between this and the straight line.

(*b*) *Faint photographic meteors.*—Photographic observations using the Super-Schmidt meteor cameras have extended the height–velocity measurements to meteors approximately five magnitudes fainter than the small camera observations. A comprehensive list of about 360 meteors has been published by Hawkins and Southworth (1958), and the variation of mean visual magnitude with velocity is given in Fig. 1 (ii). A constant colour correction of 1·5 magnitudes has been applied to convert the photographic magnitudes given by Hawkins and Southworth to visual magnitudes. A line of the correct slope for constant L is again fitted to the points. The mean magnitudes vary from $+3{\cdot}7$ for $V=13$ km/sec, to $+1{\cdot}9$ for $V=70$ km/sec. The height–log velocity relationship for these meteors is given in Fig. 2 (ii). Lines of the theoretical slopes are again shown for comparison, and it can be seen that for fast meteors the results fit quite well with a $p_{\max} \propto V^{-3}$ relationship. The slow meteors lie close to the theoretical curve when the correction term proposed by Weiss is included.

Hawkins and Southworth compared their results with a $p_{\max} \propto V^{-2}$ relation, which differs by one power of velocity from the expression used in this paper. This difference arises because Hawkins and Southworth assumed a constant meteor mass at all velocities, whereas for constant light the mass decreases by a factor of about 200 from the slowest to fastest meteors in Fig. 2.

The two series of photographic observations show that the dependence of meteor height upon velocity agrees well with theory. The variation with magnitude (or mass) can be determined from the displacements between the two magnitude–velocity curves in Fig. 1, and the height–velocity relationships in Fig. 2. The height difference is 13 km, corresponding to a difference in pressure by a factor of 6·4. The difference in magnitudes for a constant velocity is approximately 5·1, or a factor of 105 in L and I. Assuming a power law of the form

$$p_{\max} \propto L_{\max}{}^{c} \propto I_{\max}{}^{c} \qquad (12)$$

a value of $c=0{\cdot}4$ is obtained. This agrees reasonably well with the theoretical value of 0·33 (eqn. (10)). The Super-Schmidt meteors appear to vaporize rather higher than expected, and this may possibly be explained by the fragmentation of faint meteors into a number of smaller particles (Jacchia 1955).

(*c*) *The radio echo observations.*—The radio echo technique enables observations to be extended to even fainter meteors, and a series of height–velocity measurements have been made by the authors during the course of an upper atmosphere investigation. Histograms of the height distributions for different velocity groups are shown in Fig. 3. The observations were made at a frequency of

36 Mc/s, and the sensitivity of the equipment was such that the mean value of α for the echoes in a given velocity group was about 5×10^{11} electrons cm^{-1}. The experimental technique has been described in detail elsewhere (Greenhow and Hall 1960).

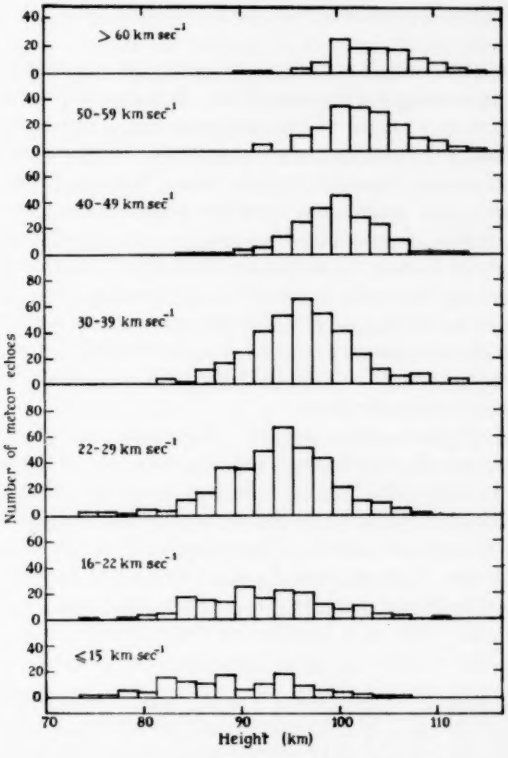


FIG. 3.—*Histograms showing the radio echo height distributions for meteors in various velocity groups.*

Various estimates of the electron densities in the trails of visual meteors have been made, and for a velocity of 60 km sec^{-1} these vary between +5 and +6 mag. for $\alpha = 10^{12}$ cm^{-1} (Greenhow and Hawkins 1952, Kaiser 1953, Davis, Greenhow and Hall 1959). A value of $\alpha = 10^{12}$ cm^{-1} for a meteor of magnitude +5·5 will therefore be adopted. $\alpha = 5 \times 10^{11}$ cm^{-1} therefore corresponds to a meteor of absolute visual magnitude +6·2. This is 4·2 magnitudes fainter than the Super-Schmidt measurements.

The expected height–velocity curve for the radio echo observations, obtained by extrapolating in magnitude from the photographic observations, is shown in Fig. 2 (iii). The measured height–velocity relation is also given (curve iv). It can be seen that there is a very large discrepancy between the observed and theoretical curves. Meteors in the very lowest velocity group, with $V = 13$ km sec^{-1}, appear to ionize at the mean height expected from theory. Meteors with a velocity of 60 km sec^{-1}, on the other hand, have a mean height 10 km below the predicted

value. In fact the mean height of these meteors is only the same as those observed photographically, although there is a factor of about 50 in mass between the two series of observations. The slope of the velocity–height relation for the radio measurements varies from that corresponding to $p_{max} \propto V^{-1}$ to $p_{max} \propto V^{-1.5}$, compared with the expected V^{-3} variation. Possible explanations for these very large departures from theory will now be considered.

4. *Discussion*

(*a*) *Density of the meteoric material.*—A meteor of constant mass and velocity vaporizes after intercepting a given mass of air. It therefore produces its maximum light and ionization at a height of constant pressure. This only applies as long as the cross-section of the meteor remains constant. If the density of the meteoric material increases, and the cross-section decreases, the meteor intercepts a smaller mass of air per unit path length, and therefore penetrates to a lower height. The variation of p_{max} with ρ_m is given by equation (1). As the observed value of p_{max} for fast radio meteors is four times greater than expected, this would require an increase in meteor density by a factor of 8. This explanation implies that faint +6 magnitude meteors do, in fact, ionize lower than expected. A second possibility is that the meteors ionize at the predicted heights, but are not detected because of an attenuation in echo amplitude. Factors which could cause such an attenuation are discussed below.

(*b*) *The effect of finite meteor velocity.*—The radio echo from a meteor trail comes essentially from the first Fresnel zone on the trail. If, however, the meteor occurs in a region where the rate of diffusion is so large that, by the time the meteor reaches the end of the zone, the ionized column has diffused to a radius $\geqslant \lambda/2\pi$, then the echo is attenuated. The magnitude of this effect depends upon the length of the first Fresnel zone $L_F = \sqrt{(R\lambda/2)}$, the meteor velocity V, the diffusion coefficient D and wavelength λ (R is the range). The variation of echo amplitude with time, as a function of the quantity

$$C = L_F/2\pi V\tau$$

has been determined by Loewenthal (1956) where $\tau = \lambda^2/16\pi^2 D$. From the observed value of the rates of decay in amplitude of meteors in the 60 km sec^{-1} group, the appropriate value of D can be determined (Herlofson 1948), giving $C = 0{\cdot}05$ at 103 km. The value of C for 60 km sec^{-1} meteors at a height of 113 km would therefore be 0·2. This would lead to a reduction of only a factor of two in amplitude. If the true mean height of 60 km sec^{-1} meteors was 113 km. then from the observed widths of meteor height distributions it would be expected that there would be ten times as many echoes from this height as from a point 10 km below the maximum. Even with a reduction of 2 in amplitude, the ratio in rates at the two heights would still be approximately 5. As the attenuation factor varies only slowly with height, large numbers of meteors influenced by finite velocity effects should still be observed at even greater heights. The observed height distribution on the other hand, cuts off completely at 113 km. Thus finite velocity effects cannot account for the large difference between the observed heights and expected heights of the radio echo observations.

(*c*) *Rapid initial diffusion.*—When individual meteor atoms vaporize from the body of a meteor, they are initially moving with forward velocities between 10 and 70 km sec^{-1}. This velocity is two orders of magnitude greater than the normal diffusive velocities for a temperature of about 300 °K. The result is that the

ions apparently diffuse at a considerably greater rate than expected, until they are slowed down by collisions with air molecules. This effect has been considered by Manning (1958) who shows that the trail radius rapidly expands to about 14 ion mean free paths, after which the normal ambipolar diffusion coefficient is predominant. The result of this rapid expansion is to reduce the initial echo amplitude. At a frequency of 36 Mc/s the effect becomes important above 110 km, but below 105 km the reduction in echo amplitude is negligible. As meteors occurring as low as 95 km have already departed considerably from the expected height, it can be concluded that rapid initial diffusion of the ions does not account for the deviations from theory.

(*d*) *Initial trail radius.*—Neither (*b*) nor (*c*) results in sufficient attenuation to account for the failure to detect echoes from +6 magnitude meteors, if they ionized at the predicted heights. A third possibility, the effect of which cannot be predicted theoretically, is that a meteor does not behave as a single body but disintegrates into a cluster of particles on entering the atmosphere. The result of this would be the formation of an ionized column with a finite initial radius r_i, and if $r_i > \lambda/2\pi$ considerable attenuation in echo amplitude would result. To reduce the echo amplitude by a factor of 10 at 36 Mc/s would require an initial radius of $1{\cdot}5\lambda/2\pi = 2$ metres. Evidence that meteors fragment into a large number of smaller particles is given by the photographic observations (McCrosky 1958), and a mechanism of this type must be considered.

(*e*) *Conclusions.*—The two probable explanations for the apparent deviations of the heights of faint meteors from the predicted heights are therefore:

(i) that +6 magnitude meteors evaporate at a lower height than suggested by theory, because of an increase in density of the meteoric material for meteors of small mass. Such an effect would appear first for meteors of high velocity, as, for a constant magnitude, the mass of a meteor is proportional to V^{-4}. This behaviour is observed;

(ii) because of the large initial radii of meteor trails, which could well be a function of height and velocity, the echo amplitudes are attenuated so that only those meteors which penetrate well below the mean height give detectable echoes.

From a study of the radio echoes obtained at a frequency of 36 Mc/s it is not possible to distinguish between these two mechanisms, although the problem could be resolved by observing with a sufficiently long radio wave-length such that $r_i < \lambda/2\pi$ at all heights.

Acknowledgments.—The research work described in this paper has been carried out at the Jodrell Bank Experimental Station, University of Manchester. The authors wish to thank Professor A. C. B. Lovell, Director of the Station, for his interest in the investigation. The radio echo observations were obtained during the course of a programme supported in part by the Air Research and Development Command, United States Air Force, under contract No. AF 61(514)-948.

Jodrell Bank Experimental Station,
Lower Withington,
Macclesfield, Cheshire:
1960 *March.*

References

Ceplecha, Z., 1958, *Bull. Astr. Inst. Czech.*, **9**, 157.
Cook, A. F., and Millman, P. M., 1955, *Ap. J.*, **121**, 250.
Davis, J., Greenhow, J. S. and Hall, J. E., 1959, *Proc. Roy. Soc. A.*, **253**, 130.
Evans, S., 1954, *M.N.*, **114**, 63.
Greenhow, J. S. and Hall, J. E., 1960, *J. Atmos. Terr. Phys.* (in publication).
Greenhow, J. S. and Hawkins, G. S., 1952, *Nature*, **170**, 355.
Hawkins, G. S. and Southworth, R. B., 1958, *Smith. Cont. Astrophys.*, **2**, 349.
Herlofson, N., 1948, *Phys. Soc. Rep. Prog. Phys.*, **11**, 444.
Jacchia, L. G., 1952, Harv. Col. Observ. & M.I.T. Center of An. Tech. Rep., 10.
Jacchia, L. G., 1955, *Ap. J.*, **121**, 521.
Kaiser, T. R., 1953, *Phil. Mag. Suppl.*, **2**, 495.
Loewenthal, M., 1956, Mass. Inst. Tech. Lincoln Lab., Tech. Rep., 132.
Manning, L. A., 1958, *J. Geophys. Res.*, **63**, 181.
McCrosky, R. E., 1958, *Astron. J.*, **63**, 97.
Öpik, E. J., 1933, *Pub. Obs. Astr. Univ. Tartu.*, **26**, (2).
Öpik, E. J., 1958, *Amer. J. Phys.*, **26**, 70.
Weiss, A. A., 1958, *Aust. J. Phys.*, **11**, 591.
Whipple, F. L., 1943, *Rev. Mod. Phys.*, **15**, 246.
Whipple, F. L., 1955, *Ap. J.*, **121**, 241.

THE IMPORTANCE OF INITIAL TRAIL RADIUS ON THE APPARENT HEIGHT AND NUMBER DISTRIBUTIONS OF METEOR ECHOES

J. S. Greenhow and J. E. Hall

(Communicated by the Director, Jodrell Bank Experimental Station, University of Manchester)

(Received 1960 March 22)

Summary

Some simultaneous radio echo observations of meteors at wavelengths of 8 m and 17 m are described. It is shown that very many more echoes from faint +6 mag. meteors are observed at 17 m than at the shorter wavelength, and this effect is attributed to an attenuation in echo amplitude due to the large initial radii of the ionized trails. The initial radius is found to increase from 1 m to 3 m between the heights of 90 and 115 km. It is estimated that radio echo equipments of moderate sensitivity detect only 1·5 per cent of +6 mag. meteors at a wavelength of 4 m, 8 per cent at 8 m, rising to 40 per cent at 17 m. The influence of this large attenuation in echo amplitude on metre wave radio echo observations of meteor height, mass, and velocity distributions is considered.

1. *Introduction*

This paper is concerned with the heights and rates of the meteors detected by radio echo methods. In a previous paper (Greenhow and Hall 1960a) it was shown that while the heights of meteors brighter than +2 magnitude vary with velocity and magnitude in the manner predicted by theory, serious departures from theory occur for the radio echo observations of +6 magnitude meteors. Although the mean heights of very slow meteors are close to the expected values, as the velocity increases, the heights become progressively lower than predicted. The discrepancy amounts to about 10 km for a 60 km/sec meteor. This result implied that either these faint meteors produce their maximum ionization in a region where the atmospheric density is up to 4 times greater than suggested by theory, or that meteors vaporize at the expected heights but for some reason the echo amplitudes are attenuated by a large factor. The suggestion was made that the required attenuation could arise if the initial radii of the meteor trails were of the order of $\lambda/2\pi$ at the observing wavelength ($\lambda = 8{\cdot}3$ m). The attenuation caused by such an effect would be much reduced at longer wavelengths, and this paper describes a two wavelength experiment designed to differentiate between the two possible explanations of the anomolous height–velocity relationship.

2. *Technique*

(*a*) *Relation between echo decay time and height.*—When the linear electron density α in a meteor trail is less than 2×10^{12} cm^{-1}, and the initial trail radius is $\ll\lambda$, then the electrons in any cross-section of the trail scatter independently and in phase. As the ionized column diffuses the echo amplitude decays exponentially with a time constant

$$\tau = \frac{\lambda^2}{(16\pi^2 D)},$$

where D is the ambipolar diffusion coefficient (Herlofson 1948). For an isothermal atmosphere D is inversely proportional to density and pressure, and both log D and log τ vary linearly with height, D and τ changing by a factor $1/e$ in a scale height H. Measurements of the echo decay time τ, which is proportional to the square of the wavelength, thus enable the height of the reflecting point to be determined (Greenhow and Neufeld 1955). In this paper an isothermal atmosphere is assumed, with a mean scale height of 7 km at a height of 100 km.

(*b*) *The echo amplitude.*—The initial echo amplitude A_0 of these 'decay' echoes is given by the expression

$$A_0 \propto (PG_T G_R \mu \lambda^3)^{1/2} \alpha \qquad (1)$$

(Lovell and Clegg 1948). P is the peak transmitter power, G_T and G_R the transmitting and receiving aerial gains and μ the pulse length (assuming optimum bandwidth).

For a finite initial trail radius r_i, where r_i is defined by the width of an assumed gaussian distribution of electron density, the initial echo amplitude is reduced to a value A_i, where

$$A_i = A_0 \exp\left[- \left(\frac{2\pi r_i}{\lambda} \right)^2 \right]. \qquad (2)$$

However, the echo amplitude should still decay with the time constant τ. When $\alpha > 2 \times 10^{12}\,\mathrm{cm}^{-1}$, on the other hand, the electron density at the centre of the trail remains greater than the critical electron density n_c for trail radii $> \lambda/2\pi$. In this case the reflection can be considered to take place at the boundary where the electron density is equal to n_c. (Greenhow 1952, Kaiser and Closs 1952). The echo amplitudes from trails of this type do not decay exponentially, but remain substantially constant for all trail radii as long as the axial electron density is greater than n_c. These echoes, which are normally only a small percentage of the total number observed and correspond to the brighter meteors, can be used to compare the sensitivities of the two radio echo equipments. Thus if an echo is observed to have the characteristic constant amplitude phase corresponding to $\alpha > 2 \times 10^{12}\,\mathrm{cm}^{-1}$, it can be assumed that no attenuation has taken place due to the finite initial radius or other effects.

(*c*) *Apparatus.*—Observations of meteor echoes have been made simultaneously at a wavelength of $\lambda_1 = 8{\cdot}3$ m, at which the height–velocity measurements were made, and at approximately twice this wavelength ($\lambda_2 = 17$ m). In order to resolve the small amplitude fluctuations which enable the meteor velocity to be determined (Davies and Ellyett 1949), pulsed transmitters operating at 600 and 300 pulses per second were used at the two wavelengths respectively. These high pulse rates also ensured that no limitations were placed on the smallest values of τ which could be measured.

The sensitivities of the equipments were equalized, so that echoes should have appeared with the same signal to noise ratio at both wavelengths. Similar aerial systems consisting of single full wave dipoles, at heights of $3\lambda_1/8$ and $3\lambda_2/8$ above the ground were used, in order to reduce variations in the comparative echo amplitude due to polar diagram differences ($G_T = G_R$ in eqn. (1)). Assuming a cosmic noise limitation in receiver sensitivity, with noise power $\epsilon \propto \lambda^{2{\cdot}5}$, the

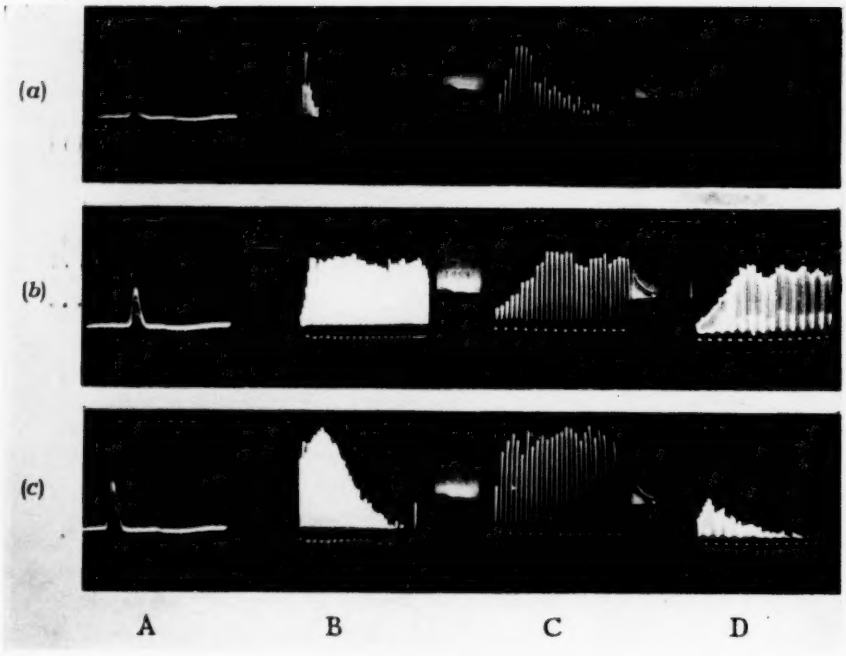


FIG. 1.—*Sections of the film records.*

(*a*) *A ' decay ' type echo at a wave-length of 17 m, not detected at 8·3 m.*
(*b*) *' Long duration ' echoes at 17 m and 8·3 m, with approximately equal amplitudes.*
(*c*) *An echo in the transition region from ' decay ' type to ' long duration ' at 17 m, which appears as a ' decay ' echo at 8·3 m.*

sensitivities can then be equalized by varying P and μ. A peak power of 10 kw and pulse lengths 30 μs were used at 8·3 m, and a similar sensitivity was achieved with $P = 2$ kw and $\mu = 100$ μs at 17 m.

The outputs of the receivers were recorded on four cathode ray tubes (Fig. 1). Tube A measured echo range at the longer wavelength. A slow time base with a duration of 0·1 sec is displayed on tubes C and D, and this enables the individual echo pulses, separated by 1·7 and 3·3 millisec, to be resolved. The 8·3 m receiver output appears as a vertical deflection on D, and the 17 m signal on C. The shortest decay times which it is possible to observe, and also the diffraction fluctuations which give the meteor velocity, can easily be measured using this 0·1 sec time base. As the values of τ at 17 m should be 4 times as long as those at 8·3 m, for coincident echoes, the decay times of meteors occurring low in the atmosphere are too long to be measured on C. For this reason a very slow time base with a duration of 1 sec is displayed together with the output of the 17 m receiver, on tube B.

The slow time bases are triggered by meteor echoes when their amplitudes rise above a preset level on either of the two equipments. Between successive triggers the cathode ray tubes are blacked out, and the film winds on in readiness for the next echo.

3. *Results*

(*a*) *Meteors brighter than* +3 *magnitude.*—In the first instance the echo amplitudes and rates from meteors brighter than about +3 mag. were compared at the two wavelengths. The electron line densities produced by these meteors are of the order 5×10^{13} cm^{-1}, giving rise to the so called 'over critically dense' trails discussed in Section 2 (*b*). These echoes can therefore be selected by the constant amplitude-time characteristic (Fig. 1 (*b*)), which distinguishes them from the echoes associated with fainter meteors, which are of the exponential decay type (Fig. 1 (*a*)). The meteors selected in this way are therefore considerably brighter than those which showed the height anomalies, and approach the magnitudes of the photographic meteors.

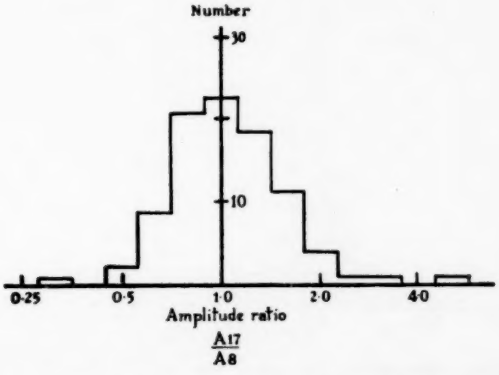


FIG. 2.—*Ratios in amplitudes of 'long-duration' echoes observed simultaneously at the two wavelengths of 8·3 m and 17 m.*
Ordinates: number of echoes.
Abscissae: ratio.

Only echoes with amplitudes greater than six times the noise level on either receiver have been included in the analysis, in order that a reasonable amplitude ratio could be measured. It was found that every echo of this type which appeared at a wavelength of 8·3 m also appeared at 17 m with about the same amplitude (Fig. 1 (*b*)). The ratios in echo amplitudes at λ_1 and λ_2 are shown by the histograms in Fig. 2. The mean ratio is 1·13, 60 per cent of the values lying between 0·7 and 1·4. This standard deviation of a factor 1·4 is readily explained by aerial polar diagram differences. Thus the amplitude ratios of the echoes are close to the theoretical value of unity, expected from the equalization of the equipment sensitivities, and it can be assumed that no echoes from +3 magnitude meteors are lost by attenuation effects.

In view of this result it can be assumed that the true amplitude ratio should be unity, and that the observed deviation is due to errors in equalizing the equipment sensitivities. The observed ratios for all echoes have therefore been corrected by this small factor.

(*b*) *Meteors of +6 magnitude.*

(i) *Comparison of echo rates at wavelengths of 17 m and 8·3 m.*—The majority of echoes observed were of the decay type, from trails with $\alpha < 2 \times 10^{12}\,\mathrm{cm}^{-1}$, and the mean magnitude of the observations was estimated to be +6·5 (Greenhow and Hawkins 1952, Davis, Greenhow and Hall 1959). When the amplitudes and rates of these echoes are compared the results are found to be very different from those for the brighter meteors. Histograms of echo rates at the two wavelengths are shown in Fig. 3, where the numbers of echoes in decay time intervals are plotted. Log decay times are plotted, and the abscissae are therefore linear in height. The actual decay times ($T_{1/2}$) measured were those to half amplitude, for which $T_{1/2} = 0{\cdot}69\,\tau$.

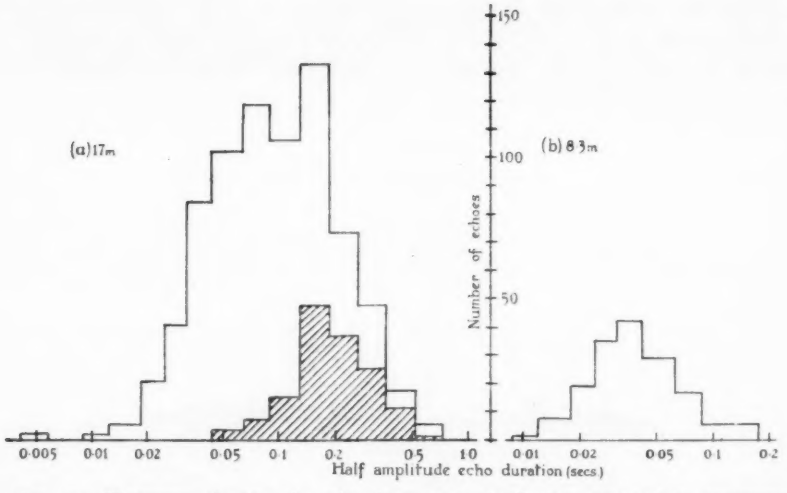


FIG. 3.—*Histograms showing the echo decay time distributions observed by equipments of equal sensitivity at wavelengths of 17 and 8·3 m. The cross hatched portion of (a) represent the echoes coincident with those of (b).*

Ordinates: number of echoes.
Abscissae: echo decay time $T_{1/2}$ (sec).

The 8·3 m histogram is shown in Fig. 3 (*b*). It can be seen that the mean value of $T_{1/2}$ is about 0·03 sec, the shortest decay times observed being about 0·01 sec. The 17 m distribution should, of course, be the same as this, but displaced by a factor of 4 in decay time to allow for the difference in wavelength. This behaviour is not observed. The cross hatched portion of Fig. 3 (*a*) are those 17 m echoes which were coincident with the 8·3 m echoes. These are displaced by approximately a factor of 4 in $T_{1/2}$, thus verifying the λ^2 duration law. However, in addition to the coincident echoes very many more were observed which did not appear at the shorter wavelength. Thus the long wave histogram still cuts off at 0·01 sec, and in view of the λ^2 variation of T, this implies that echoes are still being observed in a region where the atmospheric density is 4 times less than at the 8·3 m cut-off. The corresponding increase in height is about 10 km. The only echoes at 17 m which also appeared at 8·3 m were those with the longest decay periods, which are those of the lowest height. It is clear therefore, that many echoes from +6 magnitude meteors which should be detected at $\lambda = 8{\cdot}3$ m are not observed, and the corresponding attenuation in amplitude is strongly height-dependent. The actual ratio in numbers between the two wavelengths is 5.

(ii) *Comparison of echo rates at wavelengths of* 8·3 *m and* 4·3 *m.*—In view of the above result an attempt has been made to compare observations at 8·3 and 4·3 m, to determine whether a further reduction in rate occurs over a second factor of two in wavelength. Extensive measurements of height, velocity and decay time distributions at these two wavelengths have been made during an upper atmosphere investigation. (Evans 1954, Greenhow and Hall 1960 b). Although the two series of measurements were not coincident, both were made with equipments of similar sensitivity with aerials illuminating the same region of sky. The echo rates as a function of decay time, for similar daily observing times at the same time of the year, are shown in Fig. 4. It can be seen that once again both

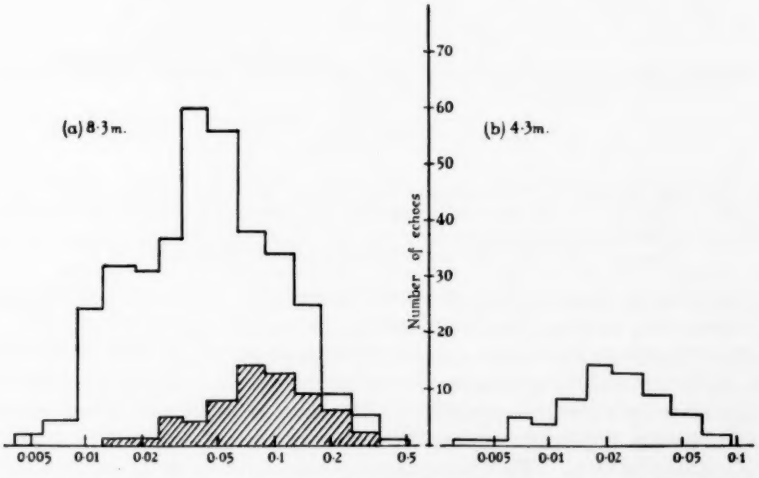


FIG. 4.—*Histograms showing the echo decay time distributions observed by equipments of equal sensitivity at wavelengths of* 8·3 *and* 4·3 *m.*
Ordinates: number of echoes.
Abscissae: echo decay time $T_{1/2}$ *(sec).*

histograms cut off at the same value of $T_{1/2}$, although a displacement of about a factor of four in decay time would be expected. The echo rate is very much higher at the longer wavelength, the ratio in numbers being approximately 6.

Although the observations were not coincident, the 4·3 m histogram has been scaled by a factor 4 in decay time, and is superimposed on the 8·3 m histogram for comparison. It can be seen that the only echoes which appear on a 4·3 m equipment are those which correspond to the longest decay times and lowest heights at $\lambda = 8{\cdot}3$ m. Thus the same effect exists between 4·3 m and 8·3 m, as was found between 17 m and 8·3 m, and only about one-thirtieth of the echoes from +6 mag. meteors which are observed at the longest wavelength would also be observed at $\lambda = 4.3$ m.

(iii) *The corrected height distribution.*—At wavelengths between 4 m and 17 m it is clear that there exist upper limits in the height distribution above which the attenuation in echo amplitude is so great that no meteors are detected. There is no reason to suppose that the true distribution has been reached even at $\lambda = 17$ m, as the upper limit of $T_{1/2} \approx 0{\cdot}01$ sec seems to be common to all wavelengths. Assuming that this cut-off in numbers is essentially an effect which is a function of the decreasing air density at increasing heights, an attempt has been made to derive the true height distribution in the following way.

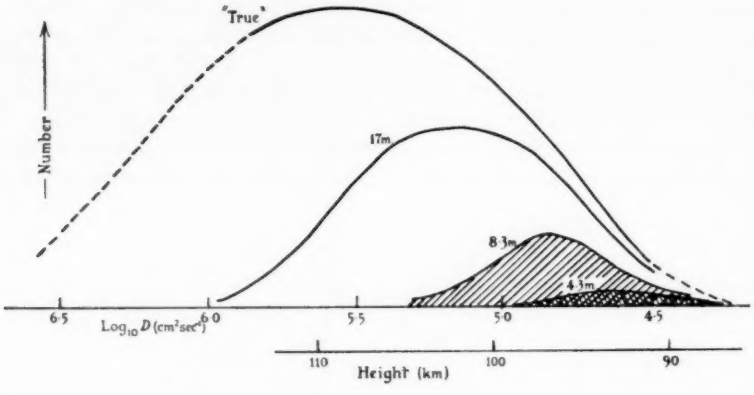


FIG. 5.—*The observed height distributions of meteors at wavelengths of* 4·3, 8·3 *and* 17 *m compared with the estimated true distribution.*
Ordinates: relative number of echoes.
Abscissae: $\log_{10}$ (*diffusion coefficient* ($cm^2\ sec^{-1}$)), *and height* (*km*)

A given decay time $T_{1/2}'$ at $\lambda = 8{\cdot}3$ m is selected. The value of $T_{1/2}$ is multiplied by a factor of 4 to give the corresponding decay time group at $\lambda = 17$ m. The numbers of meteors in the two groups are then compared, and this gives a minimum factor by which the number of echoes in the $T_{1/2}'$ group at 8·3 m must be multiplied to give the true number. It is then assumed that the attenuation factor is density dependent so that the same correction factor can be applied to the $T_{1/2}$ group at 17 m. This procedure is repeated for all decay periods, to give a new distribution at the longer wavelength. The whole operation is then repeated several times using the revised 17 m distributions, until the numbers no longer increase. In practice it was only necessary to carry out the operation

twice. As no echoes were observed at either wavelength above 115 km, no correction factor could be determined above this height. A symmetrical distribution has therefore been assumed about the point of maximum.

The results of this analysis are shown in Fig. 5, where the number-height distribution at 4·3, 8·3 and 17 m are compared with the true curve deduced in the above manner. The decay time scales have been converted to one of diffusion coefficient in order to remove the wavelength dependence of the abscissae. An approximate height scale is also shown. It can be seen that the true height distribution for +6 mag. meteors should peak at about 109 km, and some echoes should be observed as high as 130 km The maxima of the observed distributions, on the other hand, are 103 km at $\lambda = 17$ m, 97 km at $\lambda = 8{\cdot}3$ m, and 92 km at $\lambda = 4.3$ m. The percentages of the true numbers of meteors which give detectable radio echoes at the three wavelengths are 40 per cent, 8 per cent and 1·5 per cent respectively.

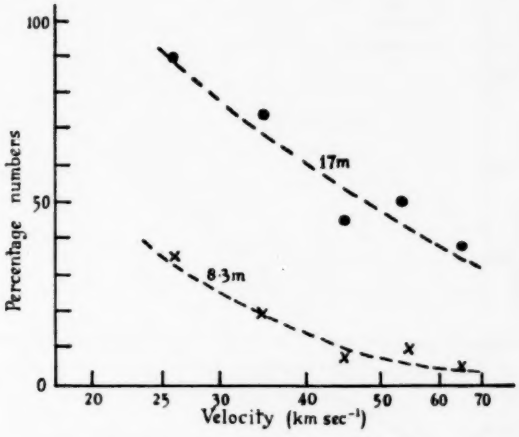


FIG. 6.—*Number of echoes observed at wavelengths of 17 m and 8·3 m, as a percentage of the estimated true number, at different velocities.*
Ordinates: percentage.
Abscissae: velocity (km/sec) (log. scale).

(iv) *The variation of observed echo rates with velocity.*—As fast meteors ionize considerably higher in the atmosphere than slow meteors, it would be expected that the percentage of trails which do not give detectable echoes would be a function of velocity. The curves in Fig. 5 apply to the total numbers of echoes observed between $23^h\,00^m$ and $06^h\,00^m$ local time irrespective of velocity. The observations have therefore been divided into velocity groups, and the ratios in numbers at the two wavelengths determined. True height distributions have also been estimated, and the percentages of the meteors giving detectable echoes at wavelengths of 8·3 and 17 m are given in Fig. 6. At 17 m the percentage falls from as high as 80 per cent at a velocity of 25 km/sec, to 30 per cent at 70 km/sec, while at 8·3 m the percentage varies from 35 per cent at 25 km/sec to only 4 per cent at 70 km/sec.

(v) *Height-velocity relationship from radio echo observations.*—The height velocity relationships for the radio echo observations of +6 mag. meteors, for

the three wavelengths, are shown for comparison with the predicted curve (Greenhow and Hall 1960a) in Fig. 7. Although absolute heights have not been measured at 17 m, the difference in heights between similar velocity groups at 8·3 m and 17 m could be determined from the differences in mean echo decay times. The 4·3 m results are due to Evans (1954). It can be seen that as the wavelength increases, the experimental curves approach the theoretical curve more closely, beginning first with the slowest meteors. At 17 m the experimental results fit the expected $p_{max} \propto V^{-3}$ curve quite well up to a velocity of 30 km/sec. This is in agreement with the results of Fig. 6, which show that most of the meteors with velocities less than this give detectable echoes. At higher velocities the observed points fall below the expected heights, as more and more echoes are lost.

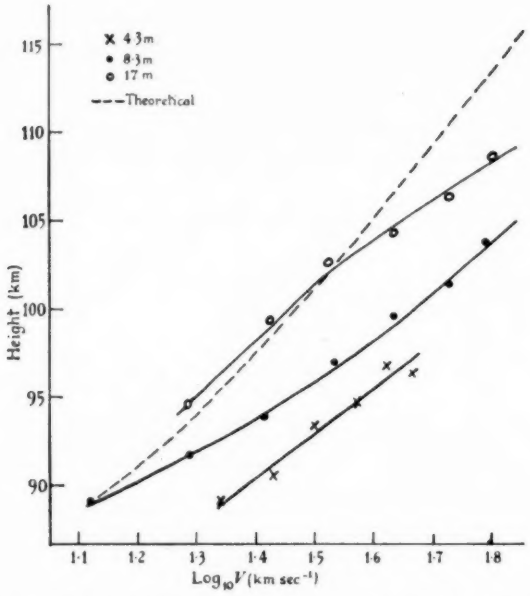


FIG. 7.—*Height–velocity relationships for meteors observed by radio echo methods, compared with the predicted curve.*
Ordinates: height (km).
Abscissae: $\log_{10}$ (velocity (km/sec)).

(*c*) *The attenuation in echo amplitude and initial trail radius*

(i) *The effect of rapid diffusion on echo amplitude.*—In the previous paper (Greenhow and Hall 1960a), the effects of rapid diffusion of the ionization were suggested as a possible cause of the anomalous height–velocity relationship, observed by radio-echo methods. This effect will now be considered quantitatively.

The section of meteor trail from which the radio echo is received is effectively the first Fresnel zone near the specular reflecting point, and in the derivation of the scattering formula it is assumed that this length of trail L_f, is formed instantaneously. If the time taken for the meteor to traverse the first zone is comparable

with τ, then before the zone is completed the ionized column near the beginning has already diffused to a radius of the order $\lambda/2\pi$. For this reason the full echo amplitude for an infinitely narrow trail is never attained. As $\tau \propto \lambda^2$, whereas $L_f \propto \lambda^{1/2}$, this effect is more important at shorter wavelengths. The maximum echo amplitude is a function of the quantity

$$C = \frac{L_f}{2\pi V \tau}$$

(Loewenthal 1956), and for a given meteor echo the appropriate value of C can be determined from the measured range, decay time and velocity.

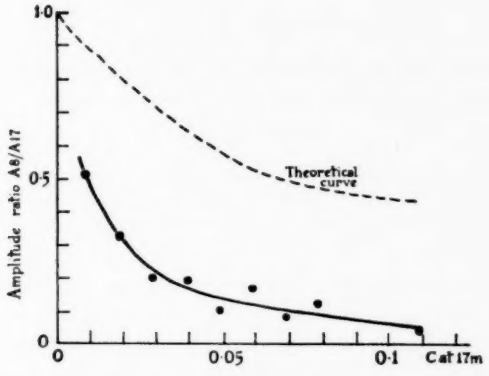


FIG. 8.—*The ratio in echo amplitudes at wavelengths of 8·3 m and 17 m, as a function of C.*
Ordinates: amplitude ratio.
Abscissae: $C = L_f/2\pi V\tau$

In order to compare the relative attenuations in echo amplitude between wavelengths of 8·3 m and 17 m, with the attenuations expected from the different values of C at the two wavelengths, the ratio in echo amplitudes for echo pairs has been plotted as a function of C in Fig. 8. Measurements have been restricted to those echoes with amplitudes greater than 15 times receiver noise at the longer wavelength, so that large ratios can be measured. It can be seen that whereas for meteors in the group with the lowest value of C, the amplitude ratio is as high as 0·5, for $C = 0{\cdot}1$ at 17 m the ratio decreases to only 0·05. The experimental measurements in Fig. 8 are also compared with the predicted curve, based on the calculation of Loewenthal. The attenuation in echo amplitude at 8·3 m, compared with that at 17 m, is very much greater than would be expected for the effect of finite meteor velocity. At $C = 0{\cdot}1$ the additional attenuation is a factor of 10, and at higher values of C at 17 m no corresponding 8·3 m echoes could even be detected.

(ii) *The variation of initial trail radius with height.*—As the effects of rapid diffusion of the trails are not sufficient to account for the observed attenuation in echo amplitudes, the results have been considered in terms of a finite initial trail radius.

Assuming that a meteor trail is formed with an initial radius r_i, then at $\lambda_1 = 8{\cdot}3$ m the initial echo amplitude given by eqn. (2) is

$$A_i(8) = A_0 \exp\left[-\left(\frac{2\pi r_i}{\lambda_1}\right)^2\right] \tag{3}$$

and at $\lambda_2 = 17$ m.

$$A_i(17) = A_0 \exp\left[-\left(\frac{2\pi r_i}{\lambda_2}\right)^2\right]. \tag{4}$$

The ratio in echo amplitude 'P' is then given by

$$P = \frac{A_i(8)}{A_i(17)} = \exp\left[-(2\pi r_i)^2\left(\frac{1}{\lambda_1^2} - \frac{1}{\lambda_2^2}\right)\right] \tag{5}$$

and as $\lambda_2 \approx 2\lambda_1$,

$$P = \exp\left[-3\left(\frac{2\pi r_i}{\lambda_2}\right)^2\right]. \tag{6}$$

The ratio in echo amplitudes at the two wavelengths therefore enables the initial radius to be determined.

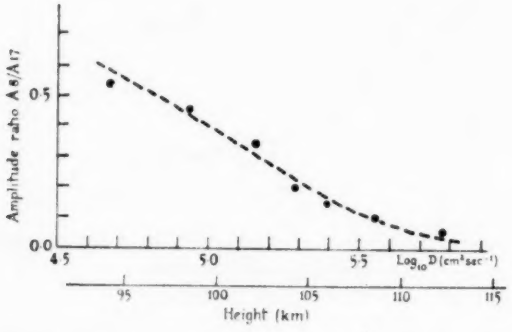


FIG. 9.—*The ratio in echo amplitudes at wavelengths of 8·3 m and 17 m as a function of diffusion coefficient and height.*
Ordinates: amplitude ratio.
Abscissae: $\log_{10}$ (Diffusion coefficient (cm^2/sec)) and height (km).

In order to investigate the height variation of r_i, P has been plotted as a function of log D (determined from the amplitude decays at λ_2), in Fig. 9. An approximate height scale is also shown. The ratio in amplitudes falls from 0·5 at a height of 95 km to only 0·05 at 112 km. The initial radii required to account for these observations have been deduced using eqn. (6), and $\log r_i$ is shown as a function of height in Fig. 10. r_i increases slowly from 1 m at a height of 90 km to 3 m at a height of 115 km. A straight line has been fitted to the experimental points, and this gives the experimental equation

$$r_i \propto \rho^{-0{\cdot}35} \tag{7}$$

for the variation of r_i with atmospheric density.

The broken curve in Fig. 9 is the theoretical curve for the variation of P with height, for an initial radius which varies in this way. There is a good fit with the observed points.

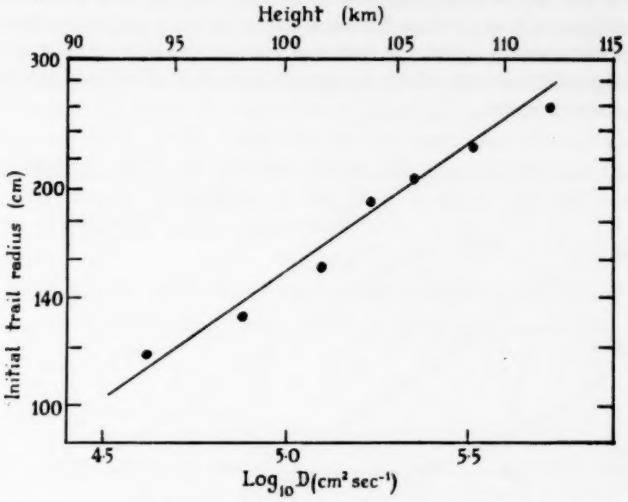


FIG. 10.—*Variation of initial trail radius with diffusion coefficient and height.*
Ordinates: initial radius (cm) (log scale).
Abscissae: Log_{10} (Diffusion coefficient (cm^2/sec)) and height (km).

(iii) *The variation of r_i with meteor velocity.*—The variation of r_i with meteor velocity has been investigated by selecting meteors in the same height range but with different velocities. The height range, 95 km to 100 km, was defined using the echo decay times at 17 m, and the ratio in echo rates at the two wavelengths were then compared. The ratio is shown as a function of velocity in Fig. 11. If P varies due to a variation of r_i with velocity, then a variation in the comparative rates at the two wavelengths would be expected. The broken line shows that variation which would be expected for $r_i \propto V$. It can be seen that the observed dependence of initial radius upon velocity is appreciably less than the first power of V.

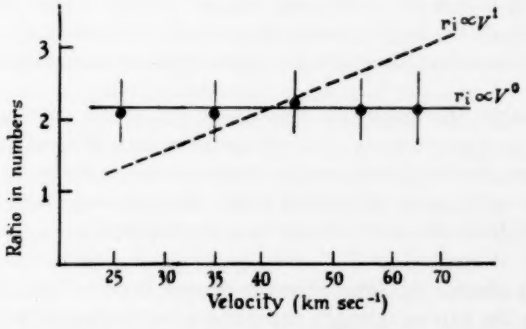


FIG. 11.—*Ratio of numbers of echoes observed at wavelengths of 17 m and 8·3 m, in the height range 95 to 100 km, as a function of velocity.*
Ordinates: ratio.
Abscissae: velocity (km/sec). Logarithmic scale.

As the initial radius is independent of meteor velocity, it is possible to use the measured values of r_i to predict the attenuation in echo amplitude for a range of radio frequencies and heights. A series of curves are shown in Fig. 12 and these show the large attenuations which are possible for V.H.F. observations at heights within the meteor region.

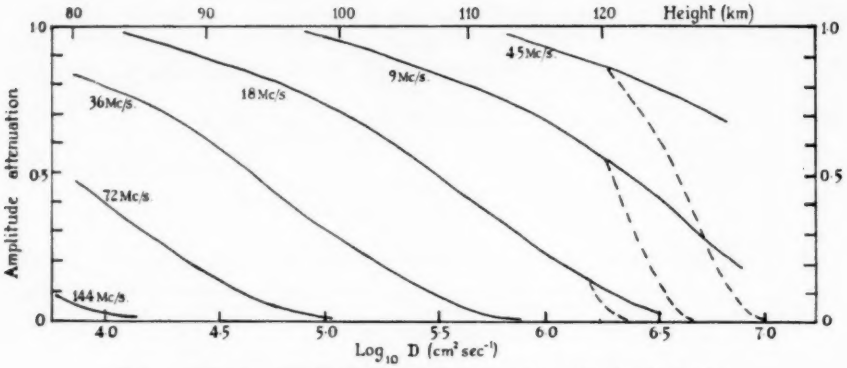


FIG. 12.—*Predicted variation of attenuation in echo amplitude with height. Curves are drawn for a number of radio frequencies.*
Ordinates: attenuation factor.
Abscissae: $\log_{10}$ (Diffusion coefficient (cm^2/sec)) and height (km).

(iv) *Determination of r_i from the amplitude variations of echoes.*—Further evidence that meteor trails have initial radii of the order of 1 to 3 m is given by comparison of the amplitude–time variations of echo pairs at 17 m and 8·3 m. Echoes are frequently observed at the longer wavelength, showing the constant amplitude phase followed by an exponential decay, which is characteristic of trails with α above the transition value of 2×10^{12} electrons/cm. An example is shown in Fig. 1 (*c*), Tube B. The 8·3 m echo, on the other hand, may appear as a simple decay echo (Tube D), and from an inspection of this echo alone, it would be concluded that the trail was underdense with $\alpha < 10^{12}$ electrons/cm. The explanation of this phenomena is that the linear electron density is greater than 2×10^{12} cm^{-1}, but that the initial trail radius is so great that the initial volume electron density on the axis is already below n_c at the shorter wavelength. Thus at 8.3 m the echo enters the exponential decay phase immediately on formation of the trail.

As an example, the amplitude–time variations of the 17 m echo in Fig. 1 (*c*) show that α is approximately 4×10^{12} electrons/cm (Kaiser 1953). In order that the volume electron density at the centre of this trail should be less than the critical density at 8·3 m, an initial trail radius of 1·8 m is required. This is of the order deduced from the attenuations in echo amplitude.

4. *Discussion*

(*a*) *The production of large diameter meteor trails.*—The initial trail radii deduced from the two wavelength experiment are compared with the ion mean free paths, determined from the measured diffusion coefficient in Table I. It can be seen that the initial radius can be as large as several hundred mean free paths at the lowest heights. The theories of meteor evaporation (Whipple 1943)

have been based on the assumption that individual meteor atoms leave the surface of the meteor, and then produce light and ionization by collision processes. If, however, the evaporation takes place in this way, the ionized trail would have an initial radius of only a few mean free paths, as the single atoms are rapidly slowed down to thermal velocities by collisions (Manning 1958). Thus in order to explain high values of r_i, it is necessary to postulate that the meteor material ablates in the form of small particles which have an appreciable range in air. Evidence that meteors fragment in this way is given by the photographic observations (Jacchia 1955, McCrosky 1958).

TABLE I

The initial radii of meteor trails, and the ion mean free path, as a function of height.

Height (km)	Initial radius (cm)	Ion mean free path (cm)
81	55	0·1
86	70	0·2
91	96	0·5
96	122	1·0
101	155	2
106	214	5
111	270	10
116	342	20
121	435	40

Recently, Hawkins and Whipple (1958), have carried out an analysis of photographic meteor trails obtained unintentionally during observations with the 48-inch Schmidt telescope at Mount Palomar. The resolving power of this instrument is sufficiently high to enable the visible trails, which were from meteors of about +2 magnitude, to be resolved. The measured radii were found to be approximately 1 metre, this value being almost independent of the position along the trails. This is the order of size required to account for the observed attenuations in the amplitudes of radio echoes.

(*b*) *The effect of finite initial trail radii on the apparent meteor mass distribution.*— An important effect of the large initial radii of meteor trails is that the radio echo observations may give very misleading measurements of number counts and height distributions. This criticism holds for observations of meteors of about +6 magnitude at wavelengths shorter than approximately 17 m, and the situation becomes worse for still fainter meteors. Estimates of $s=2$ for the exponent in the number-mass distribution $dN \propto m^{-s}dm$ have been made, where dN is the number of meteors with masses between m and $m+dm$. This value of s corresponds to an increase in the total number of meteors by 2·5 times per magnitude (Kaiser 1953). Fig. 5 shows that estimates based on the shapes of measured height distributions cannot be significant. An attempt has been made to correct the meteor echo rate counts of McKinley (1951), using the amplitude attenuations described in this paper. It is found that $s=2{\cdot}45 \pm 0{\cdot}15$ for meteors of about +6 to +8 magnitude, corresponding to an increase in total numbers by a factor of approximately 3·8 per magnitude. This is close to the value 3·4 determined photographically for meteors brighter than +2 mag. (Hawkins and Upton 1958).

In order to extend the measurements of s to fainter meteors, for example, those of +11th to +12th magnitude, which ionize a further 10 km higher than those of +6 mag., Fig. 12 shows that it would be necessary to observe at a frequency

of about 5 Mc/s. Measurements extrapolated from higher frequencies would be very uncertain, as the correction factors become so large as to be indeterminate. The situation is further complicated at heights above 120 km, as the initial radius due to the mean free path of the ions becomes significant. (Manning 1958). Above this height the initial radius would increase in proportion to the air density, and the effect of this upon the echo amplitude is shown by the broken curves in Fig. 12.

Acknowledgments

The research work described in this paper has been carried out at the Jodrell Bank Experimental Station of the University of Manchester. The work was supported in part by the Air Research and Development Command, United States Air Force, under contract No. AF 61(514)-948.

Jodrell Bank Experimental Station,
Lower Withington,
Macclesfield, Cheshire:
1960 *March.*

References

Davies, J. G., and Ellyett, C. D., 1949, *Phil. Mag.*, **40**, 614.
Davis, J., Greenhow, J. S., and Hall, J. E., 1959, *Proc. Roy. Soc. A.*, **253**, 130.
Evans, S., 1954, *M.N.*, **114**, 63.
Greenhow, J. S., 1952, *Proc. Phys. Soc. B.*, **65**, 169.
Greenhow, J. S., and Hall, J. E., 1960 a, *M.N.*, **121**, 174.
1960 b, *J. Atmos. Terr. Phys.* (in publication).
Greenhow, J. S., and Hawkins, G. S., 1952, *Nature*, **170**, 355.
Greenhow, J. S., and Neufeld, E. L., 1955, *J. Atmos. Terr. Phys.* **6**, 133.
Hawkins, G. S., and Upton, E. K. L., 1958, *Ap. J.*, **128**, 727.
Hawkins, G. S., and Whipple, F. L., 1958, *Astron. J.*, **63**, 283.
Herlofson, N., 1948, *Phys. Soc. Rep. Prog. Phys.*, **11**, 444.
Jacchia, L. G., 1955, *Ap. J.*, **121**, 521.
Kaiser, T. R., 1953, *Phil. Mag. Suppl.*, **2**, 495.
Kaiser, T. R., and Closs, R. L., 1952, *Phil. Mag.*, **43**, 1.
Loewenthal, M., 1956, M.I.T. Lincoln Lab, Tech. Rep. 132.
Lovell, A. C. B., and Clegg, J. A., 1948, *Proc. Phys. Soc.*, **60**, 491.
Manning, L. A., 1958, *J. Geophys. Res.*, **63**, 181.
McCrosky, R. E., 1958, *Astron. J.*, **63**, 97.
McKinley, D. W. R., 1951, *Can. J. Phys.*, **29**, 403.
Whipple, F. L., 1943, *Rev. Mod. Phys.*, **15**, 246.

MAGNETO-HYDROSTATICS OF STELLAR ATMOSPHERES

I. The Stability of the Axially Symmetric Case

M. J. Laird

(Communicated by H. Bondi)

(Received 1960 April 5)*

Summary

The equation governing a small perturbation in a rotating stellar atmosphere under gravity with an axially symmetric magnetic field, whose lines of force are confined to the meridional planes, is set up. A sufficient condition for stability is found to be that a certain quadratic form is positive definite.

In 1957, Hain, Lüst and Schlüter (1) made a thorough investigation of the stability of a plasma in magneto-hydrostatic equilibrium. In particular, they found a sufficient condition for stability in a perfectly conducting inviscid fluid with an axially symmetric magnetic field whose lines of force were confined to the meridional planes. In this paper, a similar method is used to study the stability of the atmosphere of a rotating star. We shall suppose that the atmosphere has the properties mentioned above, the axis of symmetry of the magnetic field and the axis of rotation being the same diameter of the star. Initially, the atmosphere is in equilibrium, the equations of which are as follows:

$$\rho\boldsymbol{\omega}\times(\boldsymbol{\omega}\times\mathbf{R})+\nabla p=\rho\mathbf{g}+\mathbf{j}\times\mathbf{H}$$
$$\operatorname{curl}\mathbf{H}=4\pi\mathbf{j}$$

giving

$$\rho\boldsymbol{\omega}\times(\boldsymbol{\omega}\times\mathbf{R})+\nabla p=\rho\mathbf{g}-\frac{1}{4\pi}\mathbf{H}\times\operatorname{curl}\mathbf{H} \tag{1}$$

where p is the pressure, ρ the mass-density, $\mathbf{g}$ the gravitational field of the star, $\boldsymbol{\omega}$ the angular velocity vector (supposed constant), $\mathbf{j}$ the current density, and $\mathbf{H}$ the magnetic field. We now suppose that this state is perturbed, so that the velocity relative to axes rotating with the star, at a point in the atmosphere, is $\mathbf{v}$, and that this is accompanied by changes in pressure, density and magnetic field of p', ρ' and $\mathbf{h}$ respectively.

The linearized equations of this perturbation are as follows:

$$\frac{\partial\mathbf{h}}{\partial t}=\operatorname{curl}(\mathbf{v}\times\mathbf{H}) \tag{2}$$

which represents the fact that the lines of force are "frozen" into the material, since it is a perfect conductor.

The equation of motion:

$$\rho\left(\frac{\partial\mathbf{v}}{\partial t}+2\boldsymbol{\omega}\times\mathbf{v}\right)=-\nabla p'+\rho'\left(\mathbf{g}-\boldsymbol{\omega}\times(\boldsymbol{\omega}\times\mathbf{R})\right)-\frac{1}{4\pi}\mathbf{H}\times\operatorname{curl}\mathbf{h}-\frac{1}{4\pi}\mathbf{h}\times\operatorname{curl}\mathbf{H}. \tag{3}$$

* Received in original form 1959 December 17.

The equation of continuity:

$$\frac{\partial \rho'}{\partial t} = -\operatorname{div} \rho \mathbf{v}. \tag{4}$$

The adiabatic equation (the conduction of heat is neglected):

$$\frac{d}{dt}(p+p') = \frac{\gamma p}{\rho}\frac{d}{dt}(\rho+\rho'). \tag{5}$$

From this last equation, and equation (4)

$$\frac{\partial p'}{\partial t} + \mathbf{v}.\nabla p = \frac{\gamma p}{\rho}\left(\mathbf{v}.\nabla\rho + \frac{\partial \rho'}{\partial t}\right) = -\gamma p \operatorname{div} \mathbf{v},$$

$$\frac{\partial p'}{\partial t} = -\operatorname{div} p\mathbf{v} - (\gamma - 1)p \operatorname{div} \mathbf{v}. \tag{6}$$

Taking $\frac{\partial}{\partial t}$ of (3) and substituting in it from (2), (4) and (6) gives the following differential equation for $\mathbf{v}$:

$$\rho\left(\frac{\partial^2 \mathbf{v}}{\partial t^2} + 2\boldsymbol{\omega} \times \frac{\partial \mathbf{v}}{\partial t}\right) = \nabla(\operatorname{div} p\mathbf{v} + (\gamma-1)p \operatorname{div} \mathbf{v}) - \mathbf{G} \operatorname{div} \rho\mathbf{v} - \frac{1}{4\pi}\mathbf{H} \times \operatorname{curl}\operatorname{curl}(\mathbf{v}\times\mathbf{H}) - \frac{1}{4\pi}\operatorname{curl}(\mathbf{v}\times\mathbf{H}) \times \operatorname{curl}\mathbf{H}, \tag{7}$$

where

$$\mathbf{G} = \mathbf{g} - \boldsymbol{\omega} \times (\boldsymbol{\omega} \times \mathbf{R}). \tag{8}$$

We now suppose that the perturbation varies with time as $e^{i\lambda t}$ so that taking a scalar product of (7) with $\mathbf{v}$ gives

$$-\rho\lambda^2 v^2 = \mathbf{v}.\nabla \operatorname{div} p\mathbf{v} + (\gamma-1)\mathbf{v}.\nabla(p \operatorname{div} \mathbf{v}) - \mathbf{v}.\mathbf{G} \operatorname{div} \rho\mathbf{v} - \frac{1}{4\pi}\mathbf{v}.\mathbf{H} \times \operatorname{curl}\operatorname{curl}(\mathbf{v}\times\mathbf{H}) - \frac{1}{4\pi}\mathbf{v}.\operatorname{curl}(\mathbf{v}\times\mathbf{H}) \times \operatorname{curl}\mathbf{H}. \tag{9}$$

A solution for which λ has a negative imaginary part indicates that the perturbation increases in magnitude exponentially in time, so that the system departs more and more from its equilibrium position at least until the linear equations become invalid, and we conclude that the position is unstable. We therefore seek to establish a criterion for λ^2 to be positive for all possible perturbations, as it is stable equilibrium configurations in which we are interested.

First of all we get rid of all the second derivatives of $\mathbf{v}$ in (9) by writing them as divergences and taking a volume integral:

$$\mathbf{v}.\nabla \operatorname{div} p\mathbf{v} = \operatorname{div}(\mathbf{v} \operatorname{div} p\mathbf{v}) - p(\operatorname{div}\mathbf{v})^2 - \mathbf{v}.\nabla p \operatorname{div}\mathbf{v}, \tag{10}$$

$$\mathbf{v}.\nabla(p \operatorname{div}\mathbf{v}) = \operatorname{div}(\mathbf{v}p \operatorname{div}\mathbf{v}) - p(\operatorname{div}\mathbf{v})^2. \tag{11}$$

$$\mathbf{v}\times\mathbf{H}.\operatorname{curl}\operatorname{curl}(\mathbf{v}\times\mathbf{H}) = -\operatorname{div}\{(\mathbf{v}\times\mathbf{H})\times\operatorname{curl}(\mathbf{v}\times\mathbf{H})\} + (\operatorname{curl}(\mathbf{v}\times\mathbf{H}))^2, \tag{12}$$

$$\mathbf{v}.\mathbf{G} \operatorname{div} \rho\mathbf{v} = -\mathbf{v}.\rho\mathbf{G} \operatorname{div}\mathbf{v} + 2\mathbf{v}.\rho\mathbf{G} \operatorname{div}\mathbf{v} + \mathbf{v}.\mathbf{G}\mathbf{v}.\nabla\rho. \tag{13}$$

Putting the force due to the magnetic field equal to $\mathbf{b}$, i.e.

$$\mathbf{b} = -\frac{1}{4\pi}\mathbf{H} \times \operatorname{curl}\mathbf{H} = \nabla p - \rho\mathbf{G} \tag{14}$$

the volume integral throughout the atmosphere of (9) is:

$$\lambda^2 \int_v \rho v^2 \, d\tau = -\int_S \left\{ \mathbf{v} \operatorname{div} p\mathbf{v} + (\gamma - 1) p\mathbf{v} \operatorname{div} \mathbf{v} + \frac{1}{4\pi} (\mathbf{v} \times \mathbf{H}) \times \operatorname{curl} (\mathbf{v} \times \mathbf{H}) \right\} . \, d\mathbf{S}$$
$$+ \int_v \left\{ \gamma p \, (\operatorname{div} \mathbf{v})^2 + \frac{1}{4\pi} (\operatorname{curl} (\mathbf{v} \times \mathbf{H}))^2 + \mathbf{v} . \mathbf{b} \operatorname{div} \mathbf{v} + 2\mathbf{v} . \rho \mathbf{G} \operatorname{div} \mathbf{v} \right.$$
$$\left. + \mathbf{v} . \mathbf{G}\mathbf{v} . \nabla \rho - \frac{1}{4\pi} \mathbf{v} . \operatorname{curl} \mathbf{H} \times \operatorname{curl} (\mathbf{v} \times \mathbf{H}) \right\} d\tau \qquad (15)$$

where S is the boundary of the volume. At the surface of the star, the lines of force in the atmosphere are anchored to those of the star itself. Since these lines are "frozen" into the material, $\mathbf{v}$ is zero there, as the perturbation is confined to the atmosphere. At the outer boundary of the atmosphere, the density vanishes, so that there is no velocity field there. The surface integral therefore vanishes. We may consider the expression left under the volume integral on the right-hand side as being a quadratic form in the components of $\mathbf{v}$ and their first derivatives. As $\int \rho v^2 \, d\tau$ is positive, if this quadratic form is positive definite, λ^2 will be positive for all possible perturbations of the equilibrium position. The next stage, therefore, consists in reducing this quadratic form as far as possible to a sum of squares. To do this, it is convenient to take components of $\operatorname{curl} (\mathbf{v} \times \mathbf{H})$ in the direction of the field and at right angles to it, thus:

$$\operatorname{curl} (\mathbf{v} \times \mathbf{H}) = \frac{1}{H^2} \{ \mathbf{HH} . \operatorname{curl} (\mathbf{v} \times \mathbf{H}) - \mathbf{H} \times (\mathbf{H} \times \operatorname{curl} (\mathbf{v} \times \mathbf{H})) \}$$
$$= \frac{\mathbf{H}}{H^2} \operatorname{div} [(\mathbf{v} \times \mathbf{H}) \times \mathbf{H}] - \frac{4\pi}{H^2} \mathbf{v} . \mathbf{b}\mathbf{H} - \frac{\mathbf{H}}{H^2} \times (\mathbf{H} \times \operatorname{curl} (\mathbf{v} \times \mathbf{H})) \qquad (16)$$

using (14). It can be shown from the symmetry of the magnetic field that the electric current $\mathbf{j}$ is toroidal, that is, perpendicular to the meridional planes, so that

$$\mathbf{H} . \operatorname{curl} \mathbf{H} = 0. \qquad (17)$$

Using also that $\operatorname{div} \mathbf{H} = 0$ and $\mathbf{H} . \mathbf{b} = 0$, Hain, Lüst and Schlüter wrote some of the terms in the quadratic form in terms of these components and $\mathbf{v}$, thus: Expanding $\mathbf{v}$ as

$$\frac{1}{H^2} \{ \mathbf{H} \times (\mathbf{v} \times \mathbf{H}) + \mathbf{v} . \mathbf{HH} \},$$
$$\mathbf{v} . \mathbf{b} \operatorname{div} \mathbf{v} = \operatorname{div} \left(\frac{\mathbf{v} . \mathbf{b} \mathbf{v} . \mathbf{HH}}{H^2} \right) - \frac{(\mathbf{v} . \mathbf{b})^2}{b^2 H^2} \mathbf{b} . \nabla H^2 + \frac{\mathbf{v} . \mathbf{b}}{H^2} \operatorname{div} (\mathbf{H} \times (\mathbf{v} \times \mathbf{H}))$$
$$- \frac{\mathbf{v} . \mathbf{H}}{H^2} \mathbf{H} . \nabla (\mathbf{v} . \mathbf{b}) \qquad (18)$$

and the first term vanishes on integration. Using

$$\operatorname{curl} \mathbf{H} = \frac{1}{H^2} \mathbf{H} \times (\operatorname{curl} \mathbf{H} \times \mathbf{H}) = \frac{4\pi}{H^2} \mathbf{H} \times \mathbf{b},$$
$$\frac{1}{4\pi} \operatorname{curl} (\mathbf{v} \times \mathbf{H}) . \mathbf{v} \times \operatorname{curl} \mathbf{H}$$
$$= \left\{ -\frac{\mathbf{v} . \mathbf{b}}{H^2} \operatorname{div} (\mathbf{H} \times (\mathbf{v} \times \mathbf{H})) - \frac{4\pi}{H^2} (\mathbf{v} . \mathbf{b})^2 - \frac{\mathbf{v} . \mathbf{H}}{H^2} \mathbf{H} . \nabla (\mathbf{v} . \mathbf{b}) - \frac{\mathbf{v} . \mathbf{H}}{H^2} \mathbf{v} . \mathbf{H} \times \operatorname{curl} \mathbf{b} \right\}. \qquad (19)$$

Substituting in (15) from (16), (18) and (19) gives the following expression for $\lambda^2 \int \rho v^2 \, d\tau$:

$$\int d\tau \left\{ \gamma p \, (\operatorname{div} \mathbf{v})^2 + \frac{1}{4\pi H^2} (\mathbf{H} \times \operatorname{curl} (\mathbf{v} \times \mathbf{H}))^2 + \frac{1}{4\pi H^2} (\operatorname{div} [\mathbf{H} \times (\mathbf{v} \times \mathbf{H})])^2 \right.$$
$$+ \frac{8\pi}{H^2} (\mathbf{v}.\mathbf{b})^2 - \frac{(\mathbf{v}.\mathbf{b})^2}{b^2 H^2} \mathbf{b}.\nabla H^2 + \frac{4}{H^2} \mathbf{v}.\mathbf{b} \operatorname{div} [\mathbf{H} \times (\mathbf{v} \times \mathbf{H})]$$
$$\left. + 2\mathbf{v}.\rho\mathbf{G} \operatorname{div} \mathbf{v} + \mathbf{v}.\mathbf{G}\mathbf{v}.\nabla\rho + \frac{\mathbf{v}.\mathbf{H}}{H^2} \mathbf{v}.\mathbf{H} \times \operatorname{curl} \mathbf{b} \right\}. \tag{20}$$

Completing squares we get:

$$\int d\tau \left\{ \gamma p \left(\operatorname{div} \mathbf{v} + \frac{\mathbf{v}.\rho\mathbf{G}}{\gamma p} \right)^2 + \frac{1}{4\pi H^2} (\mathbf{H} \times \operatorname{curl} (\mathbf{v} \times \mathbf{H}))^2 + \frac{1}{4\pi H^2} (\operatorname{div} [\mathbf{H} \times (\mathbf{v} \times \mathbf{H})] \right.$$
$$\left. + 8\pi \, \mathbf{v}.\mathbf{b})^2 \right\} + \int d\tau \left\{ -8\pi \frac{(\mathbf{v}.\mathbf{b})^2}{b^2 H^2} \left(b^2 + \frac{1}{8\pi} \mathbf{b}.\nabla H^2 \right) - \frac{\rho^2}{\gamma p} (\mathbf{v}.\mathbf{G})^2 \right.$$
$$\left. + \mathbf{v}.\mathbf{G}\mathbf{v}.\nabla\rho + \frac{\mathbf{v}.\mathbf{H}}{H^2} \mathbf{v}.\mathbf{H} \times \operatorname{curl} \mathbf{b} \right\}. \tag{21}$$

The expression under the first integral sign is a sum of squares. Therefore a sufficient condition for λ^2 to be positive, and the equilibrium to be stable, is that the expression under the second integral sign, a quadratic form in v_R and v_θ, the radial and transverse components of $\mathbf{v}$ in the meridional planes, should be positive for any possible perturbation. That is, the said quadratic form

$$-8\pi \frac{(\mathbf{v}.\mathbf{b})^2}{b^2 H^2} \left(b^2 + \frac{1}{8\pi} \mathbf{b}.\nabla H^2 \right) - \frac{\rho^2}{\gamma p} (\mathbf{v}.\mathbf{G})^2$$
$$+ \mathbf{v}.\mathbf{G}\mathbf{v}.\nabla\rho + \frac{\mathbf{v}.\mathbf{H}}{H^2} \mathbf{v}.\mathbf{H} \times \operatorname{curl} \mathbf{b} \tag{22}$$

should be positive definite.

The question now arising is how we apply this criterion. It is clear that in the first place a detailed knowledge of the equilibrium values of p, ρ and $\mathbf{H}$ is required. Other papers in this series will discuss solutions of the equilibrium equation, especially when there is no rotation. It will be shown that this equation, under certain assumptions about the structure of the atmosphere leads to a second-order differential equation in a scalar variable. In general, this does not have an exact solution, but approximate values of p, ρ and $\mathbf{H}$ at the points of a coordinate grid will be found. The coefficients in the quadratic form could then be evaluated at these points, and the test made to see if it is positive definite at all of them.

Kings College London,
W.C.2:

1960 *April* 4.

Reference

(1) Hain, K., Lüst, R., and Schlüter, A., "Zur Stabilitat eines Plasmas," *Z. Naturforschg.*, **12a**, 833, 1957.

MAGNETO-HYDROSTATICS OF STELLAR ATMOSPHERES

II. The Axially Symmetric Equilibrium Configurations

H. Bondi

(Received 1960 January 18)

Summary

It is shown how the equation of hydrostatic equilibrium may be reduced to a form suitable for calculation in the axially symmetric case, full account being taken of gas pressure, gravity, and magnetic forces. The nature of the solution is discussed, and a specially simple case is worked out in detail.

---

1. *Introduction.*—Stellar atmospheres contain regions in which gas pressure, gravitation, and magnetic forces are of comparable importance. A good deal of work has been done on the outermost regions, where magnetic fields predominate (and must therefore effectively be force free) and some work has also been done on regions where two of the three forces mentioned are of comparable importance. As a first step in the study of regions where the three forces play parts of similar importance, this paper and others in the series concern the especially simple case in which the entire configuration is both static* and axially symmetric. On the other hand we are concerned with all three fields, and thus the magnetic field vector $\mathbf{H}$, the pressure gradient grad p, and the gravitational acceleration $\mathbf{g}$ all enter our equations. The gravitational field is taken to be due to a spherical star of mass M, the self gravitation of the atmosphere being neglected. Boundary conditions will be imposed on a sphere of radius a, the 'surface' of the star where the magnetic field is supposed to be attached to the underlying denser matter, and where the pressure is supposed to be prescribed. The basic picture is that of a simple dipole magnetic field becoming distorted through the squeezing out of gas from the star, and although most of the treatment is more general, the final example illustrates this picture.

2. *Reduction of the fundamental equation.*—The relation div $\mathbf{H}=0$ may be used to express any axially symmetric magnetic field in terms of a function V (which is constant along each line of force) by

$$H_R=-\frac{1}{R^2\sin\theta}\frac{\partial V}{\partial\theta},\quad H_\theta=\frac{1}{R\sin\theta}\frac{\partial V}{\partial R},\tag{1}$$

where R, θ, ϕ are spherical polar coordinates centred on the star. The electric currents are then confined to the ϕ direction and

$$(\text{curl }\mathbf{H})_\varphi=\frac{1}{R\sin\theta}\frac{\partial^2 V}{\partial R^2}+\frac{1}{R^3}\frac{\partial}{\partial\theta}\left(\frac{1}{\sin\theta}\frac{\partial V}{\partial\theta}\right).\tag{2}$$

*The importance of such magnetohydrostatic configurations to an interpretation of the corona was first stressed by T. Gold (I.A.U. Symposium on Electromagnetic Phenomena in Cosmical Physics, Cambridge 1958, p. 275) who also examined qualitatively the situation considered in detail here.

Accordingly the fundamental equation of hydrostatic equilibrium

$$\operatorname{grad} p = -\frac{1}{4\pi}\mathbf{H}\times\operatorname{curl}\mathbf{H}+\rho\mathbf{g} \tag{3}$$

reduces to

$$\frac{\partial p}{\partial R} = -J\frac{\partial V}{\partial R} - \frac{GM\rho}{R^2} \tag{4}$$

$$\frac{1}{R}\frac{\partial p}{\partial \theta} = -J\frac{1}{R}\frac{\partial V}{\partial \theta} \tag{5}$$

where

$$J = \frac{(\operatorname{curl}\mathbf{H})_\varphi}{4\pi R\sin\theta} = \frac{1}{4\pi R^2\sin^2\theta}\left[\frac{\partial^2 V}{\partial R^2}+\frac{\sin\theta}{R^2}\frac{\partial}{\partial\theta}\left(\frac{1}{\sin\theta}\frac{\partial V}{\partial\theta}\right)\right]. \tag{6}$$

We shall now suppose that it is possible to express the density as the product of a function of the radius, a function of the pressure, and a function which is constant along each line of force. Though restrictive, this assumption is likely to be sufficiently general in view of the present limited understanding of the structure of outer stellar atmospheres. Accordingly we put

$$\frac{GM\rho}{R^2} = \frac{L(V)\,S'(R)}{P'(p)} \tag{7}$$

where S', P' denote the derivatives of functions S, P. In this formula S' and P' are supposedly known from the general structure of the atmosphere, while L is to be determined from the boundary conditions at the surface of the star, as will be shown below.

By (4)

$$\frac{\partial}{\partial R}[P(p)+L(V)S(R)] = -[JP'-SL']\frac{\partial V}{\partial R} \tag{8}$$

whereas by (5)

$$\frac{\partial}{\partial \theta}[P(p)+L(V)S(R)] = -[JP'-SL']\frac{\partial V}{\partial \theta}. \tag{9}$$

Equations (8) and (9) imply that the Jacobian

$$\frac{\partial(P+LS,\,V)}{\partial(R,\theta)} = 0$$

and thus

$$P(p) = N(V)-L(V)S(R). \tag{10}$$

With this, (8) and (9) reduce to

$$N'(V)-L'(V)S(R) = -P'(p)J. \tag{11}$$

In (10) and (11) the reduction of the fundamental equation has been achieved. The content of this pair of equations becomes clear when the boundary conditions are considered.

Taking the surface of the star as $R=a$, we suppose that there H_R, ρ and p are given as functions of θ. By (1) H_R determines V there in terms of θ except for an arbitrary constant, and by convention we put $V=0$ on $\theta=0$. Normally H_R will be taken to be on $R=a$ as given by a central dipole of moment m so that on $R=a$ $V=-m\sin^2\theta/a$. Equation (7) may now be applied at $R=a$ to determine $L(V)$ from the variation of ρ and p with θ. Finally equation (10) is used to determine $N(V)$.

In order to consider the significance of (10) and (11) in the space $R \geqslant a$, define a function F by means of the rule that if $P(p) = u$, $1/P'(p) = F'(u)$. Then (11) takes the form

$$J = \frac{1}{4\pi R^2 \sin^2\theta}\left[\frac{\partial^2 V}{\partial R^2} + \frac{\sin\theta}{R^2}\frac{\partial}{\partial\theta}\left(\frac{1}{\sin\theta}\frac{\partial V}{\partial\theta}\right)\right] = -\frac{\partial}{\partial V}F[N(V) - L(V)S(R)]. \quad (12)$$

This is the differential equation satisfied by V.

3. *Character of the solutions.*—We first consider the specially simple case when gravity is negligible. Although our derivation of (12) no longer applies, it is easily seen from (4) and (5), that p is a function of V, i.e. $p = N(V)$. The function $N(V)$ is determined by the distribution of pressure at $R = a$, and the equation corresponding to (12) becomes

$$\frac{\partial^2 V}{\partial R^2} + \frac{\sin\theta}{R^2}\frac{\partial}{\partial\theta}\left(\frac{1}{\sin\theta}\frac{\partial V}{\partial\theta}\right) = -4\pi R^2 \sin^2\theta\, N'(V). \quad (13)$$

We next study the behaviour of the field at large distances. Either the lines of force emanating from the neighbourhood of the poles close at large distances, with the field extending to infinity, or there is a bounding line of force with no field beyond it, or the field is disrupted. In the first case $V \to 0$ as $R \to \infty$. The second case corresponds to a 'sucking in' of the original dipole field by a decrease of pressure away from the poles. In what follows this case will be excluded, i.e. we shall assume that the pressure increases away from the poles.

It is immediately evident that unless $N'(V)$ tends rapidly to zero as $R \to \infty$, there will be no solution in which V tends to zero at infinity. Since the line of force $V = 0$ emanates from the poles, this cannot be a closed line of force unless $V \to 0$ as $R \to \infty$ in all directions. Hence the magnetic field will be completely disrupted round the equator unless $N'(V)$ tends to zero very rapidly with V, i.e. unless on $R = a$ the pressure gradient tends to zero very rapidly at the poles. Thus even a small excess pressure near the poles may completely disrupt the magnetic field.

In view of the substantial energy content of the magnetic field this result may cause surprise until the mobility of the gas along the lines of force is considered. If we start with constant pressure over the surface, and a dipole field, and then apply a small excess pressure around a symmetrically placed pair of circles of latitude, gas will flow from the interior of the star into the region U of space bounded by the lines of force originating along the borders of the rings of excess pressure, until this pressure has become constant throughout the whole region U. Even if the lines of force did not become deflected, the volume of U would be far greater if the rings of excess pressure were in high latitudes than near the equator, and would tend to infinity as the rings approached the poles. The deflection of the lines of force greatly amplifies this effect since the equatorial lines are shorter and thus far stiffer than the lines near the pole. Hence the volume of gas required to build up a steady state, and with it the energy used, tends to infinity as the region of excess pressure approaches the pole. Therefore a region of only slight excess pressure near the pole may totally disrupt the magnetic field.

Returning now to equation (13), consider the case when, for small V, $N'(V) \sim V^{\alpha}$. Since the basic dipole field gives $V \sim R^{-1}$, the deviation from the dipole field will become small compared with it if $\alpha > 5$. It is fairly evident that for $\alpha < 5$ there will be no solution with $V \to 0$ as $R \to \infty$ except the highly specialised

(and for our problem irrelevant ones) in which the atmospheric currents completely screen the dipole field. The case $\alpha = 5$ can also be shown to have no physically admissible solution. By equation (13) our result implies that near the poles

$$\frac{p - p_{\text{pole}}}{\sin^{12}\theta} \to 0 \quad \text{as} \quad \theta \to 0, \pi \tag{14}$$

is a necessary condition (and, if 12 is replaced by any number greater than 12, a sufficient condition) for the existence of an undisrupted field in the absence of gravity.

It is clear that gravity will tend to keep the gas down and so to prevent the disruption of the field. However, the strength of this tendency depends critically on the structure of the atmosphere. This is illustrated by the case of the completely isothermal atmosphere, i.e. one in which $p/\rho = \text{constant} = c^{-1}$ (say). By (7) this means $P(p) = \log p$, $S(R) = -GM/R$, $L(V) = c$. On $R = a$ we have $\log p = N(V) + \textit{constant}$, and equation (13) is unchanged except for the multiplication of the right hand side by the factor $\exp[N(V) + GMc/R]$. This factor is irrelevant to the convergence argument of the previous section and thus equation (14) applies as before.

Next consider the case in which the temperature varies inversely with the radius R. Let

$$\frac{p}{\rho} = \frac{GM}{nR} \tag{15}$$

where n is a constant.

Now

$$P(p) = \log p, \quad L(V) = 1, \quad S(R) = n \log(R/a)$$

and so (12) becomes

$$\frac{\partial^2 V}{\partial R^2} + \frac{\sin\theta}{R^2}\frac{\partial}{\partial\theta}\left(\frac{1}{\sin\theta}\frac{\partial V}{\partial\theta}\right) = -\frac{4\pi a^n}{R^{n-2}}\sin^2\theta\, N'(V)\, e^{N(V)}, \tag{16}$$

$N(V)$ again being equal to $\log p$ on $R = a$. It is clear now that provided only $(p - p_{\text{pole}})/\sin^2\theta$ is bounded (i.e. p is regular) at the poles and $n > 5$, the magnetic field will not be disrupted. These two examples show how delicate the balance of the problem is. If gravity is absent, or is everywhere effectively countered by a sufficiently constant temperature, the anchorage of the field near the poles can be only very slightly affected by local excess pressures without disruption of the field. If however gravity (characterized by the velocity of escape) wins sufficiently over temperature (characterized by the velocity of sound) at large distances, no regular pressure variation at the surface can burst the lines of force. It must, however, be emphatically stated that the type of disruption here discussed implies infinite energy sources in the surface since infinite volumes of gas must have passed through the surface to effect the disruption. If instead of the mathematically more convenient specification of pressure at the surface the available energy had been prescribed (as is physically more correct) we should merely have seen that only a limited range of pressure distributions results from any arbitrary regular distribution of available energy. Only strictly static situations have been considered. The non-existence of a static solution may only imply slow motions over limited regions and should not be taken as suggesting explosions or other major motions.

4. *Special forms of the equations.*—In view of our lack of understanding of processes in the corona, we have no accurate knowledge of the functions P and S.

It seems therefore in place to suggest a few special forms, in addition to (13) and (16), which would seem to deserve particular attention as being possible approximations to the actual state of affairs in the corona.

(i) Isothermal along lines of force: This is reasonable since the heat conductivity along a line of force greatly exceeds that across it. By (7) this implies

$$S(R)=-\frac{GM}{R},\ P(p)=\log p,\ L(V)=\frac{\mu}{\mathfrak{R}T}\qquad\left\{\begin{array}{c}\mathfrak{R}=\text{gas constant}\\ \mu=\text{mean molecular weight}\end{array}\right\}\tag{17}$$

In this case a plausible boundary condition is $p=constant=p_0$ on $R=a$.

Thus
$$N(V)=\log p_0-\frac{GM}{a}L(V)$$

$$\frac{\partial^2 V}{\partial R^2}+\frac{\sin\theta}{R^2}\frac{\partial}{\partial\theta}\left(\frac{1}{\sin\theta}\frac{\partial V}{\partial\theta}\right)=-4\pi R^2\sin^2\theta\, p_0\frac{\partial}{\partial V}\exp\left[-GML(V)\left(\frac{1}{a}-\frac{1}{R}\right)\right]$$
$$=4\pi R^2\sin^2\theta\, p_0\left(\frac{1}{a}-\frac{1}{R}\right)GML'(V)\exp\left[-GML(V)\left(\frac{1}{a}-\frac{1}{R}\right)\right].\tag{18}$$

Evidently conditions (14) apply to the temperature now.

(ii) Adiabatic along lines of force: This too is reasonable since mass motions can occur along but not across lines of force. Since $\gamma=5/3$ we now have

$$P(p)=\frac{5}{2}p^{2/5},\quad S(R)=-\frac{GM}{R},\quad L(V)=\frac{\rho}{p^{3/5}}.\tag{19}$$

If again we suppose for simplicity that the gas pressure is constant on $R=a$

$$\frac{5}{2}p_0^{2/5}-\frac{GM}{a}L(V)=N(V)$$

$$\frac{\partial^2 V}{\partial R^2}+\frac{\sin\theta}{R^2}\frac{\partial}{\partial\theta}\left(\frac{1}{\sin\theta}\frac{\partial V}{\partial\theta}\right)=\frac{\partial}{\partial V}\left[p_0^{2/5}-\frac{2}{5}L(V)\left(\frac{GM}{a}-\frac{GM}{R}\right)\right]^{5/2}$$
$$=-L'(V)\left(\frac{GM}{a}-\frac{GM}{R}\right)\left[p_0^{2/5}-\frac{2}{5}L(V)\left(\frac{GM}{a}-\frac{GM}{R}\right)\right]^{3/2}.\tag{20}$$

Conditions (14) now apply to $\rho p^{-3/5}$, i.e. effectively to the entropy.

(iii) Discontinuity in the pressure: A discontinuity in the pressure implies by (3) a discontinuity in **H**. It is readily seen from this equation that the sum of the magnetic and gas pressures, i.e. $p+\mathbf{H}^2/8\pi$, must be continuous. The simplest case occurs if gravity is neglected and the pressure is constant except for this discontinuity. The field is then by (14) sufficiently well anchored in the polar zones to contain at least some discontinuity of pressure. We shall suppose the pressure to be higher at the equator than at the pole. Then in the meridian plane the pressure discontinuity will be a common line of force separating two regions in each of which the left hand side of (11) and therefore J vanishes, i.e. there are no currents. In fact the equation $J=0$ is equivalent to $\nabla^2 W=0$ where

$$V=R\sin\theta\frac{\partial W}{\partial\theta}.\tag{21}$$

On the critical line of force **H** is discontinuous in accordance with the continuity of $p+\mathbf{H}^2/8\pi$, and, on $R=a$, V and hence W is given as usual. It is seen

easily that the discontinuity must be tangential at $R=a$ since H_R is continuous but $\mathbf{H}$ is discontinuous.

The critical line of force may be considered in a hydro-dynamical analogue as a membrane (or vortex sheet) separating two regions of irrotational flow of an ideal incompressible fluid with dipole normal velocities at $R=a$ but with a pressure discontinuity.

5. *A Special Case.*—We suppose, as in (15),

$$\frac{p}{\rho}=\frac{\mathfrak{R}T}{\mu}=\frac{GM}{nR}$$

and assume that $n>5$.

Furthermore we take at the surface

$$p=p_0+(p_1-p_0)\sin^2\theta \quad (p_1>p_0) \tag{22}$$

so that p_0 is the polar and p_1 the equatorial pressure. Taking the normal surface field to be that due to a dipole, i.e.

$$V=-\frac{m}{a}\sin^2\theta$$

we have

$$e^{N(V)}=p_0-\frac{a}{m}(p_1-p_0)V \tag{23}$$

and so (16) becomes

$$\frac{\partial^2 V}{\partial R^2}+\frac{\sin\theta}{R^2}\frac{\partial}{\partial\theta}\left(\frac{1}{\sin\theta}\frac{\partial V}{\partial\theta}\right)=\frac{4\pi a^{n+1}(p_1-p_0)}{mR^{n-2}}\sin^2\theta. \tag{24}$$

The solution satisfying the boundary condition ($V=-m\sin^2\theta/a$) and behaving correctly at infinity is

$$V=-m\sin^2\theta\left[\frac{1+\kappa}{R}-\kappa\frac{a^{n-5}}{R^{n-4}}\right] \tag{25}$$

where

$$\kappa=\frac{4\pi(p_1-p_0)a^6}{(n-2)(n-5)m^2}=\frac{16\pi}{(n-5)(n-2)}\frac{p_1-p_0}{\mathbf{H}^2_{\text{pole}}}. \tag{26}$$

Defining now a critical value of κ by $\kappa_c=1/(n-5)$ and with it a critical pressure difference $(p_1-p_0)_c=(n-2)\,\mathbf{H}_{\text{pole}}{}^2/16\pi$, it is easily seen that for $\kappa<\kappa_c$ the lines of force at the surface lean towards the equator, for $\kappa=\kappa_c$ they are radial, and for $\kappa>\kappa_c$ they lean polewards and become radial only at $R/a=[(n-4)\kappa/(\kappa+1)]^{1/(n-5)}$. For definiteness we now take $n=10$, corresponding, in the case of the Sun, to a temperature of about one million degrees at $R=a$. Then $\kappa_c=0{\cdot}2$, and the lines of force are shown in Fig. 1 for various typical values of κ.

It is readily seen that

$$p=\left(\frac{a}{R}\right)^{10}\left\{p_0+(p_1-p_0)\sin^2\theta\left[(1+\kappa)\frac{a}{R}-\kappa\left(\frac{a}{R}\right)^6\right]\right\}. \tag{27}$$

The pressure distribution therefore depends on the additional parameter $(p_1-p_0)/p_0$.

Although this examination of the problem has concentrated entirely on static aspects, the lines of force shown in Fig. 1 for $\kappa>0{\cdot}2$ are strongly suggestive of an

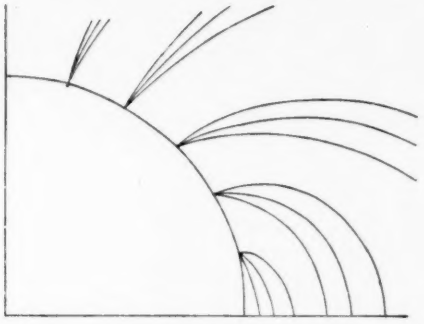


FIG. 1a.—*Lines of force emerging from the surface of a star for various pressures* ($\kappa=0, 0{\cdot}1, 0{\cdot}2$). *The innermost lines of force are those of the dipole field* ($\kappa=0$) *the outermost ones those of the field of limiting stability* ($\kappa=0{\cdot}2$) *which emerge radially from the star.*

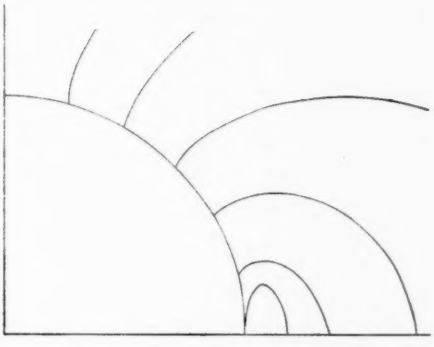


FIG. 1b.—*Lines of force in the unstable case* ($\kappa=0{\cdot}3$). *Note the closed line of force near the equator.*

unstable situation*. It is well known from many subjects that the onset of instability occurs with a change in the static pattern. It may therefore be surmised, until better information is available, that the patterns for $\kappa>0{\cdot}2$ are unstable, and so that the greatest possible pressure difference with the law (15) for $n=10$ is $p_1-p_0=5\mathbf{H}_{\text{pole}}^2/2\pi$. The existence of a neutral line is in any case believed to be connected with the onset of an electrical instability. The vertical tangent to the magnetic lines of force at the surface (such as is frequently indicated by solar prominences) therefore corresponds to the end of the stable regime.

The author wishes to record his debt to Professor T. Gold for suggesting the problem and for preliminary discussions, and to a referee of the Society for suggesting substantial improvements in the paper.

King's College London,
W.C.2:
1960 *January* 16.

* Since in this case some lines of force do not meet the surface of the star, the functions $L(V)$ and $N(V)$ are physically undefined on these lines of forces.

MAGNETO-HYDROSTATICS OF STELLAR ATMOSPHERES
III. The Axially Symmetric Equilibrium Configurations (*Continued*)

M. J. Laird

(Communicated by H. Bondi)

(Received 1960 April 5)*

Summary

In Paper II, Bondi showed how the fundamental equation may be reduced to a form suitable for calculation in the axially symmetric case under certain assumptions. In the first part of this paper a more general method of effecting this reduction is given, with an example. In the second part, this method is extended to the case when the star is rotating, under more restricted assumptions about the structure of the atmosphere.

---

1. *The non-rotating case.*—In Paper II (1), hereafter referred to as (II), Bondi showed that the fundamental equation

$$\nabla p = +\rho \mathbf{g} - \frac{1}{4\pi}\mathbf{H} \times \operatorname{curl} \mathbf{H} \tag{1.1}$$

may be written in the form

$$\nabla p = -\rho \nabla \Omega - J \nabla V \tag{1.2}$$

using the notation of (II), and the additional relationship

$$\mathbf{g} = -\nabla \Omega \quad : \quad \Omega = -\frac{GM}{R}. \tag{1.3}$$

As this yields but two scalar equations, we need a further relationship between the variables to progress. We shall consider two types of relationship. The one considered by Bondi in (II) is a special case of the first of these: (*a*) The density ρ may be expressed as a function of p and V multiplied by a function of R; and (*b*) ρ may be expressed as a function of p and R.

(*a*) *ρ may be expressed as a function of p and V multiplied by a function of R.* Let

$$\rho = f(p,V)h(R) \tag{1.4}$$

and let the function $S(R)$ be defined by

$$\nabla S(R) = h(R)\nabla \Omega(R). \tag{1.5}$$

Suppose that p is not a function of V. (If p is a function of V, then (1.2) shows that ∇V is radial, and hence both p and V are functions of R alone.) We can then use p and V as co-ordinates, instead of R and θ. Considering S as a function of p and V, (1.2) becomes:

$$\nabla p = -f(p,V)\left(\frac{\partial S}{\partial p}\nabla p + \frac{\partial S}{\partial V}\nabla V\right) - J\nabla V. \tag{1.6}$$

*Received in original form 1960 February 12.

Equating the coefficients of ∇p and ∇V separately to zero gives:

$$\frac{\partial S}{\partial p} = -\frac{1}{f(p,V)} \tag{1.7}$$

$$J = -f(p,V)\frac{\partial S}{\partial V}. \tag{1.8}$$

Integrating (1.7) we get:

$$S = -F(V) - \int \frac{dp}{f(p,V)} \tag{1.9}$$

where F is an arbitrary function of V whose form is determined by the variation of p and V with θ on the surface of the star, as explained in (II).

Equation (1.8) now gives:

$$J = f(p,V)\left\{F'(V) + \frac{\partial}{\partial V}\int \frac{dp}{f(p,V)}\right\}. \tag{1.10}$$

In (II), Bondi supposed that the function $f(p, V)$ was separable into the product of a function of p and a function of V. We shall now show that the equations derived in (II) may also be derived from equations (1.9) and (1.10) above.

Using the notation of (II) we put

$$f(p,V) = \frac{L(V)}{P'(p)} \quad ; \quad N(V) = -L(V)F(V).$$

Substituting for these gives:

$$S(R) = (N(V) - P(p))/L(V)$$

$$J = \frac{L}{P'}\left\{-\frac{N'}{L} + \frac{NL'}{L^2} - \frac{PL'}{L^2}\right\}$$

$$-P'(p)J = N'(V) - L'(V)S(R)$$

as was obtained in (II).

(*b*) *ρ may be expressed as a function of p and R.*—Supposing that p is not a function of Ω alone, we can use these variables as co-ordinates instead of R and θ. Considering V as a function of p and Ω, (1.2) becomes:

$$\nabla p = -\rho\nabla\Omega - J\left(\frac{\partial V}{\partial \Omega}\nabla\Omega + \frac{\partial V}{\partial p}\nabla p\right). \tag{1.11}$$

Equating the coefficients of ∇p and $\nabla\Omega$ separately to zero:

$$1 = -J\frac{\partial V}{\partial p} \tag{1.12}$$

$$\rho(p,\Omega) = -J\frac{\partial V}{\partial \Omega}. \tag{1.13}$$

Eliminating J between these equations gives:

$$\frac{\partial V}{\partial \Omega} - \rho\frac{\partial V}{\partial p} = \theta. \tag{1.14}$$

Let $u(p, \Omega) =$ constant be the solution of

$$d\Omega = -\frac{dp}{\rho(p,\Omega)}. \tag{1.15}$$

Then the solution of (1.14) is $V=F(u)$, where F is an arbitrary function whose form is determined by the variation of p and V with θ on the surface of the star. We can now use $V=F(u)$ to eliminate p from either (1.12) or (1.13), giving a second order partial differential equation for V of the form

$$J=G(V,R). \tag{1.16}$$

Purely as an example to illustrate the method, consider the case when ρ is written as follows:

$$\rho=\frac{nR}{GM}(p-\lambda V)$$

where λ is a constant, and $n>5$.

We suppose that at the surface $R=a$

$$p=p_0+(p_1-p_0)\sin^2\theta$$

where p_0 is the polar, and p_1 the equatorial pressure, and that the normal surface field is that due to a dipole, i.e.

$$V=-\frac{m}{a}\sin^2\theta.$$

As

$$\Omega=-\frac{GM}{R}\ (1.5)\text{ gives }\frac{dS}{dR}=\frac{n}{R}$$

i.e.

$$S=n\log R$$

so that, using (1.9),

$$S=n\log R=-F(V)-\log(p-\lambda V).$$

Applying the boundary conditions at $R=a$

$$e^{-F(V)}=a^n\left(p_0-\frac{a}{m}(p_1-p_0)V-\lambda V\right)$$

(1.10) now gives

$$J=R^{-n}e^{-F(V)}F'(V)+(p-\lambda V)\frac{\partial}{\partial V}\log(p-\lambda V).$$

The differential equation to be satisfied by V is therefore

$$\frac{\partial^2V}{\partial R^2}+\frac{\sin\theta}{R^2}\frac{\partial}{\partial\theta}\left(\frac{1}{\sin\theta}\frac{\partial V}{\partial\theta}\right)=4\pi R^2\sin^2\theta\frac{a^n}{R^n}\left[\frac{a}{m}(p_1-p_0)+\lambda\right]-4\pi\lambda R^2\sin^2\theta.$$

A solution of this equation satisfying the boundary condition

$$V=-\frac{m}{a}\sin^2\theta$$

on the sphere $R=a$ is:

$$V=-\frac{m}{a}\sin^2\theta\left(\frac{Aa}{R}+\frac{BR^2}{a^2}-\alpha\frac{a^{n-4}}{R^{n-4}}-\beta\frac{R^4}{a^4}\right)$$

where

$$\alpha=4\pi\frac{a^5\left[\frac{a}{m}(p_1-p_0)+\lambda\right]}{(n-2)(n-5)m}$$

$$\beta=-\frac{2\pi\lambda a^5}{5m}$$

and A, B, α and β are connected by the relationship

$$A+B-\alpha-\beta=1.$$

The magnetic field from the star is then entirely contained within a sphere defined by $V=0$. Outside this sphere we have $V=0$ everywhere. The additional condition

$$\frac{\partial V}{\partial R}=0$$

on this sphere must also be applied in order to avoid an impossible singularity on the axis*. Suppose that this sphere is of radius qa. Then these conditions give:

$$Aq^{n-5}+Bq^{n-2}-\alpha-\beta q^n=0$$

$$-Aq^{n-5}+2Bq^{n-2}+(n-4)\alpha-4\beta q^n=0.$$

This pair of equations, with the previous one, then enable us to determine A, B and q in terms of the other constants, so that V fulfils all of the boundary conditions.

2. *The rotating case.*—In a perfectly conducting fluid, the lines of force of the magnetic field behave as though they are "frozen" into the material. Because of the vast distances involved, it can be shown that the assumption of the "freezing in" of the lines of force in a stellar atmosphere is a fair one to make. Thus a star communicates its rotation to its atmosphere. In the case of the Sun, the centrifugal force is equal to the gravitational force at a distance of about thirty solar radii in the equatorial plane. As the Sun's atmosphere is known to extend at least this far, the centrifugal force assumes considerable importance. Dungey (**2**) gives a qualitative discussion of this problem. We shall use the following results:

If the rotation is steady, then the angular velocity is constant along a line of force (Ferraro's law of isorotation).

If further the rotation is non-uniform (as in the Sun), the field is symmetric about the axis of rotation. Therefore we may write the angular velocity vector $\boldsymbol{\omega}$ in the form $\omega\mathbf{k}$, where ω is a function of V alone, and $\mathbf{k}$ is a unit vector in the direction of the common axis of the rotation and the magnetic field. Our fundamental equation now reads:

$$\rho\omega^2\mathbf{k}\times(\mathbf{k}\times\mathbf{R})=-\nabla p-\rho\nabla\Omega-J\nabla V. \tag{2.1}$$

As

$$\omega^2\mathbf{k}\times(\mathbf{k}\times\mathbf{R})=-\frac{\omega^2}{2}\nabla R^2\sin^2\theta=-\nabla\frac{\omega^2R^2\sin^2\theta}{2}+\frac{R^2\sin^2\theta}{2}\nabla\omega^2 \tag{2.2}$$

and as ω is a function of V alone, so that

$$\nabla\omega^2=2\omega\frac{d\omega}{dV}\nabla V. \tag{2.3}$$

Defining the functions Γ and Φ by

$$\Gamma=\Omega-\frac{\omega^2R^2\sin^2\theta}{2}\ ;\ \Phi=J+\rho R^2\sin^2\theta\,\omega\frac{d\omega}{dV} \tag{2.4}$$

we get

$$\nabla p=-\rho\nabla\Gamma-\Phi\nabla V. \tag{2.5}$$

* I must thank the referee for pointing this out.

If we now assume that ρ may be expressed as a function of p and V, we can proceed as in 1 (*a*). Using p and V as co-ordinates, and considering Γ as a function of these two variables, we get a pair of equations corresponding to (1.9) and (1.10):

$$\Gamma = -F(V) - \int \frac{dp}{\rho(p,V)} \tag{2.6}$$

$$\Phi = \rho \left\{ F'(V) + \frac{\partial}{\partial V} \int \frac{dp}{\rho(p,V)} \right\} \tag{2.7}$$

where F is an arbitrary function of V whose form is determined by the variation of p, V and ω with θ on the surface of the star. We may use the first equation to eliminate p from the second, and so get a second order partial differential equation for V in the form:

$$J = G(V, \Gamma). \tag{2.8}$$

If we suppose further that $\rho(p,V)$ is separable, and may be written in the form

$$\rho = \frac{L(V)}{P'(p)} \tag{2.9}$$

putting $N(V) = -L(V)F(V)$ and substituting in (2.6) and (2.7) we get a pair of equations corresponding to those derived by Bondi in (II) for the non-rotating case, viz.

$$\Gamma = \{N(V) - P(p)\}/L(V) \tag{2.10}$$

$$-P'(p)\Phi = N'(V) - L'(V)\Gamma. \tag{2.11}$$

In particular, we may use these equations to consider two of the special forms of atmospheric structure suggested in (II) as being possible approximations to the actual state of affairs in the solar corona. These are: (*a*) isothermal along lines of force; and (*b*) adiabatic along lines of force.

In conclusion, I should like to thank Professor Bondi for suggesting the foregoing work, and for his advice during the writing of this paper.

King's College London,
W.C.2:
1960 *April* 4.

References

(1) Bondi, H., *M.N.*, **121**, 201, 1960.
(2) Dungey, J. W., *Cosmic Electro-dynamics*, Cambridge University Press, 1958.

ELECTRON IMPACT EXCITATION OF POSITIVE IONS: APPLICATION TO Ca^+ $4s$–$4p$ AND $3d$–$4p$

H. van Regemorter

(Communicated by M. J. Seaton)

(Received 1959 December 16)

Summary

A review is given of different approximate methods for computing cross sections for transitions in positive ions produced by electron impact. It is shown that it is necessary to take into account the distortion by the ion field, and, for strong optically allowed transitions, one must allow for strong coupling effects.

Application to the $4s-4p$ and $4p-3d$ transitions in singly ionized calcium shows that the neglect of strong coupling effects overestimates the cross sections. However, our results, computed with a Coulomb distorted wave approximation, with allowance for strong coupling effects, are much bigger than those of Jefferies, the reason being that he considered only one value ($l=1$) of the angular momentum of the incident electron, whereas the main contribution is shown to occur for $l=4$.

The Bethe approximation, with allowance for Coulomb distortion, implies in fact two different approximations: the weak coupling approximation and the " long range approximation ". It is possible to make corrections for strong coupling effects and, for strong allowed transitions such as $4s-4p$ and $4p-3d$ in Ca^+, this method gives good results in a very simple way.

---

1. *Introduction.*—The study of the $4s-4p$ and $4p-3d$ collision cross-sections in singly ionized calcium is interesting from two points of view. Firstly, it requires the development of a suitable approximate theory for strong optically allowed transitions in an ion, produced by electron impact. It is necessary to take into account the distortion by the Coulomb field and to allow for strong coupling effects. In the present paper we shall discuss different approximate methods.

Cross sections for $Ca^+ 4s-4p$ have been calculated by Jefferies (**8**) using a weak coupling approximation, but the work of Seaton (**11**) for strong allowed transitions in Na and H suggests that weak coupling approximations will overestimate the Ca^+ cross sections. In a later paper, Seaton (**12**) has shown that allowance for strong coupling may be made on using the weak coupling approximation to calculate the reactance matrix **R** and then deducing the transmission matrix **T** from the exact relation between **T** and **R**. In this method, conservation conditions are satisfied exactly. It is of interest to apply the method, with allowance for the Coulomb distortion, to the case of collisions between ions and electrons.

Comparison with more accurate calculations shows that it is possible to allow for the strong coupling effects in the Bethe approximation and to obtain good results in a very simple way.

On the other hand, information about the relative population of the $4s$, $4p$ and $3d$ levels is required in many astrophysical problems.

In stellar atmospheres, excitation and deactivation by electron impacts are far from negligible. Particularly, when thermodynamical equilibrium does not occur, as in the solar chromosphere, one has to know accurate cross sections for collisionally induced transitions. Excitation of the $4p$ levels gives rise to the H and K lines, for which we have many observations. These lines have abnormal intensities in many particular stars, for example in the metallic line stars and in the sub-dwarfs.

In other respects, the radiation in the K line in solar flocculi cannot be explained by radiative processes only and the emission in the K_2 line is determined by collision between Ca^+ ions and electrons. Emission lines at the centres of the H and K lines of Ca^+ are also observed in the spectral types G, K and M.

2. *The different approximations of a collision problem.*—Consider an atom having one electron moving in the field

$$V(r_1) = -\frac{Z_p(r_1)}{r_1} \tag{1}$$

where r_1 is the coordinate of the atomic electron and Z_p the effective charge of the core (nucleus screened by the inner electrons). If $H(\mathbf{r}_1) = -(\hbar^2/2m)\nabla_1^2 + V(r_1)$, the atomic electron wave functions are solutions of

$$H(\mathbf{r}_1)\Psi_{nLM}(\mathbf{r}_1) = E_{nL}\Psi_{nLM}(\mathbf{r}_1). \tag{2}$$

In a collision problem, an incident electron characterised by its initial relative velocity $\mathbf{v}$ and by its angular momentum lm is passing in the neighbourhood of the atom, and the entire system—atom and electron—initially in the state $\alpha = nLMklm$, goes to the state $\alpha' = n'L'M'k'l'm'$ where $\mathbf{k} = m_e\mathbf{v}/\hbar$ is the wave vector of the incident electron (m_e being the electron mass).

Exact wave functions of the two particle system are solutions of

$$(H - E)\Psi_\alpha(\mathbf{r}_1, \mathbf{r}_2) = 0. \tag{3}$$

The total Hamiltonian is given by

$$H = H(\mathbf{r}_1) + H(\mathbf{r}_2) + \frac{1}{r_{12}} \tag{4}$$

where $H(\mathbf{r}) = -\frac{1}{2}\nabla^2 + V(r)$ and where we use atomic units ($m_e = e = \hbar = 1$). The total energy, which is conserved during the collision, is given by

$$E = E_{nL} + \tfrac{1}{2}k^2 = E_{n'L'} + \tfrac{1}{2}k'^2; \tag{5}$$

when atomic units are used, k^2 is numerically equal to the electron kinetic energy in units of 13·60 eV. The cross section for collisionally induced transitions $nL \to n'L'$ is given by the expression

$$Q[nL \to n'L'] = \frac{\Omega(nL, n'L')}{\omega_{nL}\, k^2} \text{ in units of } \pi a_0^2 = 8{\cdot}806 \times 10^{-17}\,\text{cm}^2 \tag{6}$$

where the dimensionless parameter $\Omega(nL, n'L')$ is the collision strength and ω_{nL} is the statistical weight of the initial atomic level. With consistent neglect of spin variables, $\omega_{nL} = (2L + 1)$. Neglecting all exchange effects we can write

$$\Psi(nLMklm|\mathbf{r}_1, \mathbf{r}_2) = \psi(nLM|\mathbf{r}_1)\phi(klm|\mathbf{r}_2) \tag{7}$$

when the coupling is weak. The exact solution of the problem requires exact wave functions of the atomic electron and of the incident electron. In practice we use Hartree–Fock wave functions for the atomic electron.

2.1. *The Born I approximation.*—If we can neglect the atomic field $V(r_2)$ in evaluating the approximate wave function we have

$$(H'-E)\Psi(nLM\mathbf{k}|\mathbf{r}_1,\mathbf{r}_2)=0 \tag{8}$$

with $H'=H(\mathbf{r}_1)-\frac{1}{2}\nabla_2^2$. In such a case

$$\Psi(nLM\mathbf{k}|\mathbf{r}_1,\mathbf{r}_2)=\psi(nLM|\mathbf{r}_1)\phi(\mathbf{k}|\mathbf{r}_2) \tag{9}$$

where $\phi(\mathbf{k}|\mathbf{r}_2)=e^{i\mathbf{k}.\mathbf{r}_2}$ represents a plane wave. Supposing all interactions to be weak one obtains

$$\Omega(nL,n'L')=\frac{kk'}{4\pi^3}\sum_{MM'}\int|(nLM\mathbf{k}|H-E|n'L'M'\mathbf{k}')|^2d\omega \tag{10}$$

where

$$(H-E)\Psi(nLM\mathbf{k}|\mathbf{r}_1,\mathbf{r}_2)=\left(V(r_2)+\frac{1}{r_{12}}\right)\Psi(nLM\mathbf{k}|\mathbf{r}_1,\mathbf{r}_2) \tag{11}$$

and where the integral is over the solid angle in the direction of $\mathbf{k}$ (or $\mathbf{k}'$).

Partial wave expansions

The expansion of the plane wave can be shown to be (9, p. 22)

$$e^{i\mathbf{k}.\mathbf{r}}=k^{-\frac{1}{2}}r^{-1}\sum_l(2l+1)P_l(\hat{\mathbf{r}}.\hat{\mathbf{k}})i^lF_{kl}(r) \tag{12}$$

with (5, p. 54)

$$P_l(\hat{\mathbf{r}}.\hat{\mathbf{k}})=\sum_m\frac{4\pi}{(2l+1)}Y_{lm}(\hat{\mathbf{r}})Y^*_{lm}(\hat{\mathbf{k}}). \tag{13}$$

Our notation is that $\hat{\mathbf{r}}$ is a unit vector specifying the direction of $\mathbf{r}$, $\hat{\mathbf{r}}.\hat{\mathbf{k}}=\cos\beta$ where β is the angle between $\mathbf{r}$ and $\mathbf{k}$, P_l is a Legendre polynomial and Y_{lm} is a normalized spherical harmonic. The radial functions

$$F_{kl}(r)=\left(\frac{\pi r}{2}\right)^{\frac{1}{2}}J_{l+\frac{1}{2}}(kr) \tag{14}$$

have asymptotic form

$$F_{kl}(r)\underset{r\to\infty}{\sim}k^{-\frac{1}{2}}\sin(kr-\tfrac{1}{2}l\pi). \tag{15}$$

Thus we obtain

$$\Omega(nL,n'L')=\sum_{MlmM'l'm'}|T(nLMlm,n'L'l'm')|^2 \tag{16}$$

with

$$T(\alpha,\alpha')=-2iL(\alpha,\alpha') \tag{17}$$

and

$$L(\alpha,\alpha')=-2(\alpha|H-E|\alpha') \tag{18}$$

where

$$\Psi(\alpha|\mathbf{r}_1,\mathbf{r}_2)=\Psi(nLMlm|\mathbf{r}_1,\mathbf{r}_2)=\psi(nLM|\mathbf{r}_1)Y_{lm}(\hat{\mathbf{r}}_2)\frac{1}{r_2}F_{kl}(r_2). \tag{19}$$

Coupled angular momentum representation

The total angular momentum being conserved during the collision, it is convenient to express the wave functions in terms of $\mathbf{L}^T=\mathbf{L}+\mathbf{l}$ and $\mathbf{M}^T=\mathbf{M}+\mathbf{m}$. In such a representation

$$\Psi(nLlL^TM^T)=\sum_{Mm}C^{LlL^T}_{MmM^T}\Psi(nLMlm) \tag{20}$$

where

$$C^{abc}_{\alpha\beta\gamma}$$

is a vector-coupling coefficient. The transmission matrix $\mathbf{T}$ is diagonal in $L^T M^T$ and is independent of M^T:

$$T(nLlL^T M^T, n'L'l'L^{T'} M^{T'}) = T(nLlL^T, n'L'l'L^T)\delta(L^T, L^{T'})\delta(M^T, M^{T'}). \quad (21)$$

Parity conservation requires that $T = 0$ unless

$$(-1)^{L+l} = (-1)^{L'+l'}. \quad (22)$$

We obtain

$$\Omega(nL, n'L') = \sum_{lL^T} (2L^T + 1)|T(nLlL^T, n'L'l'L^T)|^2. \quad (23)$$

Reciprocity and conservation conditions

The reciprocity relations require that $\mathbf{T}$ should be symmetric and the conservation relations require that the scattering matrix $\mathbf{S} = \mathbf{1} - \mathbf{T}$ should be unitary ($\mathbf{1}$ being the unit matrix). The conservation condition may be expressed in the form (13)

$$|\mathbf{1} - \mathbf{T}|^2 = \mathbf{1}. \quad (24)$$

The Born I approximation will not satisfy this condition but the departures will not be serious if

$$\mathbf{T} \ll \mathbf{1}. \quad (25)$$

Then, if calculation with the Born I approximation give $\mathbf{T}$ (Born I) $\ll \mathbf{1}$, the initial assumption of weak interactions will be a good approximation, and we can use the Born I method to compute the cross section Q.

2.2. *The Born II approximation.*—For elastic scattering, $T = 1 - e^{2i\eta}$ where η is the phase (9). Putting $R = \tan \eta$ we have

$$T = -\frac{2iR}{1 - iR}. \quad (26)$$

One can show (12) that for inelastic collisions, one obtains a relation of the same kind between two matrices:

$$\mathbf{T} = -\frac{2i\mathbf{R}}{1 - i\mathbf{R}}. \quad (27)$$

$\mathbf{T}$ calculated from this relation will satisfy the conservation condition (24) if $\mathbf{R}$ is Hermitian ($\mathbf{R} = \mathbf{R}^+$).

In the Born approximation for $\mathbf{R}$, we have

$$R_{\alpha\alpha'}(\text{Born}) = L(\alpha, \alpha') \quad (28)$$

where $L(\alpha, \alpha')$ defined by (18) is calculated using plane wave function (19). The approximation Born II for $\mathbf{T}$ will be

$$\mathbf{T}\ (\text{Born II}) = -\frac{2i\mathbf{R}(\text{Born I})}{1 - i\mathbf{R}(\text{Born I})} \quad (29)$$

and in this case the conservation condition (24) will be satisfied, even if the interactions are not weak.

Using exactly soluble models, Seaton (11) has shown that the Born II approximation is much superior to Born I, when interactions become strong—which is the case for strong allowed transitions such as $2s - 2p$ in H and $3s - 3p$ in Na.

2.3. *The Coulomb–Born approximation.*—For large distances, the interaction potential is

$$v(r_2) = -z/r_2 \quad \text{with} \quad z = Z_p(\infty) - 1. \quad (30)$$

For collisions with neutral atoms $Z_p(\infty) = 1$ and $z = 0$ but for collisions with positive ions $z > 0$.

In the case of slow collisions with positive ions, it is essential to take the long range Coulomb distortion into account in the calculation of the approximate wave functions. In the Coulomb–Born approximation we use Coulomb functions in place of plane waves. The approximation CB I and CB II are similar to Born I and Born II but, in place of plane waves, Coulomb waves are used. We have

$$T(\alpha,\alpha')_{\mathrm{CBI}} = -2iL(\alpha,\alpha') \tag{31}$$

$$\mathbf{T}(\mathrm{CB\,II}) = -\frac{2i\mathbf{R}(\mathrm{CB\,I})}{1-i\mathbf{R}(\mathrm{CB\,I})} \tag{32}$$

where $L(\alpha,\alpha')$ is defined by (18) and Ψ_α by (19) but now $F_{kl}(r_2)$ is the solution of the radial equation

$$\left[\frac{d^2}{dr^2} - \frac{l(l+1)}{r^2} + \frac{2z}{r} + k^2\right] F_{kl}(r) = 0 \tag{33}$$

with asymptotic form

$$F_{kl}(r\to\infty)\to k^{-1/2}\sin\left[kr - \tfrac{1}{2}l\pi + (z/k)\ln(2kr) + \arg\Gamma(l+1-z/k)\right]. \tag{34}$$

For the case of an attractive field between a singly ionized atom and an electron, $z=1$, and one has

$$F_{kl}(r) = \left(\frac{\pi}{2}\right)^{1/2}\frac{2^{l+1}}{(2l+1)!}\left[\frac{A_{kl}}{1-e^{-2\pi/k}}\right]^{1/2} r^{l+1}\,e^{ikr}F[l+1-i/k,\ 2l+2,\ -2ikr] \tag{35}$$

with

$$A_{kl} = (1+k^2l^2)(1+k^2(l-1)^2)\ldots(1+k^2)$$
$$A_{k0} = 1$$

and where F is the confluent hypergeometric function.

For l very small, one has to take into account the exact expression for $Z_p(r_2)$, but it may be shown that it is a good approximation to put $z(r_2)=z(\infty)$ for $l>1$. When l is greater than 1, contributions to the integral (18) from values of r greater than the atomic radius are much more important than those from values of r less than the atomic radius. We shall see later on that, for the case of Ca^+, the biggest contribution to Ω is obtained for $l=4$.

Thus, for optically allowed transitions in singly ionized atoms produced by electrons, one can use expression (35) for the radial Coulomb function. In the CB approximation for inelastic collisions, we have finite values of Ω at the threshold, which is a consequence of the long range attractive Coulomb field, whereas, in the Born approximation, we would always obtain $\Omega=0$ at threshold. Thus, for collision with positive ions, the Born approximation does not give the correct threshold variation of Ω with the energy of the incident particle.

Multipole expansions

For collisions between electrons and singly ionized atoms from (11), (18) and (30) one obtains

$$L(\alpha,\alpha') = -2\left(\alpha\left|\frac{1}{r_{12}} + V(r_2)\right|\alpha'\right). \tag{36}$$

For inelastic collisions, one has

$$L(\alpha,\alpha') = -2\left(\alpha\left|\frac{1}{r_{12}}\right|\alpha'\right) \text{ with } \alpha\neq\alpha'. \tag{37}$$

Using the expansion (5, p. 174)

$$\frac{1}{r_{12}} = \sum_{\lambda} P_{\lambda}(\hat{\mathbf{r}}_1 . \hat{\mathbf{r}}_2)\gamma_{\lambda}(r_1 r_2) \tag{38}$$

with

$$\gamma_{\lambda}(r_1 r_2) = \begin{cases} r_1^{\lambda}/r_2^{\lambda+1} \text{ for } r_1 < r_2 \\ r_2^{\lambda}/r_1^{\lambda+1} \text{ for } r_1 > r_2 \end{cases} \tag{39}$$

and

$$P_{\lambda}(\hat{\mathbf{r}}_1 . \hat{\mathbf{r}}_2) = \frac{4\pi}{2\lambda+1} \sum_{\mu} Y_{lm}(\hat{\mathbf{r}}_1) Y_{\lambda\mu}^{*}(\hat{\mathbf{r}}_2) \tag{40}$$

it is possible to expand the integral (37) in terms of λ. In practice, only a few values of λ give a finite contribution to $L(\alpha, \alpha')$. In the total angular momentum representation, one has to evaluate terms of the form

$$\left(nLlL^T \left| \frac{1}{r_{12}} \right| n'L'l'L^T \right). \tag{41}$$

It may be shown that the only non-zero contributions are those for which

$$|L - L'| \leqslant \lambda \leqslant L + L' \tag{42}$$

$$|l - l'| \leqslant \lambda \leqslant l + l' \tag{43}$$

$$[L + L' + \lambda] \text{ and } [l + l' + \lambda] \text{ both even.} \tag{44}$$

It follows that only the dipole terms ($\lambda = 1$) occur for the $s - p$ transition. For the $p - d$ transition we have $\lambda = 1$ and 3 but may expect that the dipole term gives the dominant contribution.

2.4. *The Bethe approximation.*—The Bethe approximation is usually considered as a simplification of the Born approximation valid at high energies. The usual procedure is to retain only the leading terms of an expansion of (10) in power of $\mathbf{K} = \mathbf{k} - \mathbf{k}'$. This is therefore valid for small deflections.

An alternative derivation of the Bethe approximation is to replace

$$\frac{1}{r_{12}},$$

given by (38), by its form for $r_2 > r_1$

$$\frac{1}{r_{12}} = \sum_{\lambda} P_{\lambda}(\hat{\mathbf{r}}_1 . \hat{\mathbf{r}}_2) r_1^{\lambda}/r_2^{\lambda+1}. \tag{45}$$

This is valid for distant encounters, which, of course, again implies small deflections. The approximation (45) may be used with Coulomb wave functions; this approximation has been used extensively in nuclear excitation problems. With an attractive Coulomb field the contribution from large values of l at low energies is much greater than for the case of no Coulomb field (plane waves). At fairly low energies the approximation (45) for the interaction is much better with Coulomb waves than with plane waves.

One can, of course, use the Bethe approximation either in the weak coupling case (referred to as CB'I) or in the strong coupling case (referred to as CB'II). We shall discuss later on in detail the validity of these two approximations.

3. *The R matrix in the Coulomb–Born approximation.*—Using the coupled angular momentum representation ($\alpha = nLlM^TL^T$) the wave functions are given by (20) and the terms of the R matrix by

$$R_{\alpha\alpha'} = -2\left(\alpha \left| \frac{1}{r_{12}} \right| \alpha' \right) \text{ for } \alpha \neq \alpha' \tag{46}$$

If $\frac{1}{r_{12}}$ is decomposed into the multipole expansion (38) one can decompose the preceding integral into two parts, the angular integral and the radial integral. One can express $R_{\alpha\alpha'}$ in the form

$$R_{\alpha\alpha'} = -2\sum_{\lambda} f_{\lambda}\mathscr{R}_{\lambda} \tag{47}$$

with

$$f_{\lambda} = \frac{4\pi}{2\lambda+1}\sum_{MM'm'm_{\mu}} C^{LlL^T}_{MmM^T} C^{L'l'L^T}_{M'm'M^T}\int Y^*_{LM}Y_{\lambda\mu}Y_{L'M'}\,d\mathbf{P}\int Y^*_{lm}Y_{\lambda\mu}Y_{l'm'}\,d\mathbf{P}_2 \tag{48}$$

and

$$\mathscr{R}_{\lambda} = \int_0^{\infty} F_{kl}(r_2)F_{k'l'}(r_2)\int_0^{\infty}\gamma_{\lambda}(r_1r_2)P_{nL}(r_1)P_{n'L'}(r_1)\,dr_1\,dr_2. \tag{49}$$

It follows from the theory of the spherical harmonics and of the Racah coefficients that:

$$f_{\lambda} = (-1)^{L+L'-L^T}(2\lambda-1)^{-1}[(2l+1)(2l'+1)(2L+1)(2L'+1)]^{1/2} C^{ll'\lambda}_{000}C^{LL'\lambda}_{000}W[LlL'l';L^T\lambda]. \tag{50}$$

Tables of the coefficients $f_{\lambda}(LlL'l'L^T)$ are given by Percival and Seaton (**10**).

In the radial term $\mathscr{R}_{\lambda}$, the Coulomb functions $F_{kl}(r_2)$ are given by (35) and the functions $P_{nL}(r_1)$ are the atomic radial function. For the Ca^+ $4s$, $4p$ and $3d$ levels the radial functions $P_{nL}(r_1)$ are given by Hartree (**7**). If

$$y_{\lambda}(r_2) = \int_0^{\infty}\gamma_{\lambda}(r_1,r_2)P_{nL}(r_1)P_{nL'}(r_1)\,dr, \tag{51}$$

using the expansion (39) of γ_{λ} we can write

$$y_{\lambda}(r_2) = \frac{1}{r_2^{\lambda+1}}\int_0^{r_2} P_{nL}P_{n'L'}r_1^{\lambda}\,dr_1 + r_2^{\lambda}\int_{r_2}^{\infty} P_{nL}P_{n'L'}\frac{1}{r_1^{\lambda+1}}\,dr_1. \tag{52}$$

It is convenient to decompose y_{λ} in two parts, one of them representative of the long range interactions, the other decreasing steeply with r_2.

$$y_{\lambda}(r_2) = \frac{1}{r_2^{\lambda+1}}\int_0^{\infty} P_{nL}P_{n'L'}r_1^{\lambda}\,dr_1 + z_{\lambda}[nLn'L'|r_2] \tag{53}$$

with

$$z_{\lambda} = r_2^{\lambda}\int_{r_2}^{\infty} P_{nL}P_{n'L'}\frac{1}{r_1^{\lambda+1}}\,dr_1 - \frac{1}{r_2^{\lambda+1}}\int_{r_2}^{\infty} P_{nL}P_{n'L'}r_1^{\lambda}\,dr_1. \tag{54}$$

Thus, one can express $\mathscr{R}_{\lambda}$ as a sum of two integrals; the first may be evaluated analytically; the second, which depends on detailed properties of the atom, converges rapidly.

$$\mathscr{R}_{\lambda} = \int_0^{\infty} P_{nL}P_{n'L'}r_1^{\lambda}\,dr_1 \times \int_0^{\infty} F_{kl}F_{k'l'}\frac{dr_2}{r_2^{\lambda+1}} + \int_0^{\infty} z_{\lambda}(nLn'L'|r_2)F_{kl}F_{k'l'}\,dr_2 \tag{55}$$

where one recognizes the matrix elements of the 2λ-pole line strength $\mathscr{S}_{\lambda}(nL,n'L')$.

For the dipole case ($\lambda=1$) one finds

$$R_{\alpha\alpha'} = -2f_1\mathscr{R}_1 \tag{56}$$

with

$$\mathscr{R}_1 = a_1[nL,n'L']\int_0^{\infty} F_{kl}F_{k'l'}\frac{dr_2}{r_2^2} + \int_0^{\infty} z_1F_{kl}F_{k'l'}\,dr_2 \tag{57}$$

and

$$a_1[nL,n'L'] = \int_0^{\infty} P_{nL}(r_1)P_{n'L'}(r_1)r_1\,dr_1. \tag{58}$$

The evaluation of the Coulomb integrals of the type $I_\lambda = \int_0^\infty F_{kl} F_{k'l'} r^{-\lambda-1}\, dr$ has been studied in several works (**1**, **2**, **3**). Because of the very slow convergence of the dipole integral ($\lambda = 1$) it is very useful to have an analytical expression for I_1. In Appendix I we give the evaluation of I_1 by a method which converges quickly for most of the cases of interest.

It would be very useful to tabulate the Coulomb integral I_1 as a function of k, k', l, l' in order to be able to compute very quickly the cross section for any transition. With a method very similar to ours, Burgess (**4**) is programming the calculation of the Coulomb integral I_1. It is not possible to derive an analytical expression of the second integral in (57). These integrals were calculated numerically by Mr Lawson and Mrs Lawson who describe the method used in Appendix II.

4. *Evaluation of the cross sections for* $4s - 4p$ *and* $4p - 3d$ *transitions in* Ca^+.— With the CB II approximation, the **T** matrix is given by (32). For the $4p - 3d$ transition, we consider only the dipole contribution. Taking into account all the possible couplings between the different levels $4s$, $4p$ and $3d$, and using the parity conservation condition (22) one has to evaluate a $q \times q$ matrix.

The terms of the **R** matrix, given by (56) are denoted by

$$\begin{aligned}
\alpha &= R[4sl, 4pl-1] \\
\beta &= R[4sl, 4pl+1] \\
\gamma &= R[4pl-1, 3dl] \\
\delta &= R[4pl-1, 3dl-2] \\
\epsilon &= R[4pl+1, 3dl] \\
\zeta &= R[4pl+1, 3dl+2] \\
\nu &= R[4pl, 3dl+1] \\
\mu &= R[4pl, 3dl-1].
\end{aligned}$$

One finds, for the elements of the **T** matrix for the $4s - 4p$ transition, taking into account the coupling with the $3d$ level,

$$T[4sl, 4pl-1] = \frac{2i\alpha(1+\zeta^2+\epsilon^2) - 2i\beta\gamma\epsilon}{D} \tag{59}$$

$$T[4sl, 4pl+1] = \frac{2i\beta(1+\delta^2+\gamma^2) - 2i\beta\gamma\epsilon}{D} \tag{60}$$

with $$\begin{aligned} D = 1 &+ \alpha^2 + \beta^2 + \gamma^2 + \delta^2 + \epsilon^2 + \zeta^2 + \delta^2\zeta^2 + \alpha^2\zeta^2 + \gamma^2\zeta^2 + \beta^2\gamma^2 + \epsilon^2\delta^2 \\ &+ \alpha^2\epsilon^2 + \beta^2\delta^2 - 2\alpha\beta\gamma\epsilon. \end{aligned}$$

If one neglects the coupling with $3d$, one obtains

$$T(4sl, 4pl-1) = \frac{2i\alpha}{1+\alpha^2+\beta^2} \tag{61}$$

$$T(4sl, 4pl+1) = \frac{2i\beta}{1+\alpha^2+\beta^2}. \tag{62}$$

One obtains similar expressions for the **T** matrix elements for the $4p - 3d$ transition.

For the transitions from the $4s$ level, one has $L = 0$ and $L^T = l$. For the transitions from $4p$, one has $L = 1$ and $L_T = l \pm 1$ or l. For Ω given by

$$\Omega(nL, n'L') = \sum_{lL^T} (2L^T+1) |T(nLkl|L^T, n'L'k'l'L^T)|^2 \tag{63}$$

one obtains the results given in Tables I and II, where also are given the values of $\Omega_l = \sum_{l' L^T} (2L^T+1)|T|^2$ to show the contributions to Ω from the different values of l. For the $4s-4p$ transition, we give the results taking into account the $3d$ level (CB II, 3 levels) and neglecting the coupling with $3d$ (CB II, 2 levels). One notes that the coupling with the $3d$ level does not affect Ω very much. Values of Ω_l and Ω are given for 2 values of k'^2 representing the kinetic energy of the electron when the atom is in the $4p$ state. From the results of Tables I and II we see that the strong coupling effects are important and that when corrections are made for strong coupling the collision strengths are much smaller than those obtained in the weak coupling approximation CB I, for which we have $T_{\alpha\alpha'} = -2iR_{\alpha\alpha'}$.

With the CB II approximation, the cross sections given by (6) in units of πa_0^2 are

$$Q[4s \to 4p] = \begin{cases} 58{\cdot}2 & \text{for} \quad k'^2 = 0{\cdot}0 \\ 52{\cdot}0 & \text{for} \quad k'^2 = 0{\cdot}1 \end{cases} \tag{64}$$

$$Q[3d \to 4p] = \begin{cases} 58{\cdot}2_5 & \text{for} \quad k'^2 = 0{\cdot}0 \\ 31{\cdot}4_5 & \text{for} \quad k'^2 = 0{\cdot}1 \end{cases}. \tag{65}$$

For transitions between two levels defined by $nLSJ$ where $\mathbf{S}$ is the spin vector and $\mathbf{J} = \mathbf{L} + \mathbf{S}$ the total angular momentum, we have

$$\Omega(nLS, n'L'S) = \tfrac{1}{2} \sum_{S^T L^T} (2L^T+1)(2S^T+1)|T|^2 \text{ with } S_T = 0 \text{ or } 1 \tag{66}$$

$$= 2\Omega(nL, n'L'). \tag{67}$$

Consequently, with $S = \frac{1}{2}$,

$$Q[nLS \to n'L'S] = Q[nL \to n'L']. \tag{68}$$

On the other hand, if one assumes LS coupling, it is possible to show that, for transition from an s state ($L = 0$)

$$Q[nLSJ \to n'L'S'J'] = \frac{2J'+1}{(2S'+1)(2L'+1)} Q[nLS \to n'L'S']. \tag{69}$$

This gives for the transitions corresponding to the H and K lines

$$Q[4s^2 S_{1/2} \to 4p^2 P_{1/2}] = \tfrac{1}{3} Q[4s \to 4p] = \begin{cases} 19,39 & \text{for} \quad k'^2 = 0 \\ 17,35 & \text{for} \quad k'^2 = 0, 1 \end{cases} \tag{70}$$

$$Q[4s^2 S_{1/2} \to 4p^2 P_{3/2}] = \tfrac{2}{3} Q[4s \to 4p] = \begin{cases} 38,78 & \text{for} \quad k'^2 = 0 \\ 34,70 & \text{for} \quad k'^2 = 0, 1. \end{cases} \tag{71}$$

These cross sections are much bigger than those given by J. T. Jefferies (**8**). In his paper, Jefferies uses the weak coupling approximation CB I, but only takes account of the contribution Ω_l for $l = 1$ to the cross section Q. Our results show that this simplification is not correct because of the important contribution of large values of l, the maximum of Ω_l being for $l = 4$. Jefferies finds $Q[S_{1/2} \to P_{1/2}] = 6{\cdot}6$ and $Q[S_{1/2} \to P_{3/2}] = 9{\cdot}9$. We must note that the ratio of those two cross sections is not correct: as we saw just above, because of (68) and (69) the ratio should be equal to the ratio of the statistical weights

$$\frac{2J_1'+1}{2J_2'+1} \text{ with } J_1' = \tfrac{1}{2} \text{ and } J_2' = \tfrac{3}{2}, \text{ equal to } \tfrac{1}{2}.$$

5. *The Bethe approximation.*—In the Bethe approximation, one assumes both that the coupling is weak and that the long range interaction ($r_2 > r_1$) is much more important than the short range interaction. This approximation using (31) and (45) is referred to as CB′ I.

We obtain

$$T_{\alpha\alpha'} = 2iR_{\alpha\alpha'} \tag{72}$$

and

$$R_{\alpha\alpha'}(\text{Bethe}) = -2f_1\mathscr{R}_1 \tag{73}$$

with

$$\mathscr{R}_1\ (\text{Bethe}) = a_1[nL, n'L'] \int_0^\infty F_{kl} F_{k'l'} \frac{dr}{r^2}. \tag{74}$$

This gives for the collision strength

$$\Omega_{ll'}^{\text{CB'I}}(nL, n'L') = \sum_{L^T} (2L^T+1)|T|^2 = \sum_{L^T} (2L^T+1) f_1^2 16 [a_1(nL, n'L')]^2 \left| \int_0^\infty F_{kl} F_{k'l'} \frac{dr}{r^2} \right|^2 \tag{75}$$

Using (50) and relations between the Racah coefficients W we may write

$$3 \sum_{L^T} (2L^T+1)[f_1(LlL'l'L_T)]^2 = \frac{(2L+1)(2L'+1)}{2} C_{L'1L} \frac{(2l+1)(2l'+1)}{2} C_{l'1l} \tag{76}$$

with

$$C_{l'1l} = \int_{-1}^{+1} P_{l'}(t) P_1(t) P_l(t)\, dt$$

where the $P_l(t)$ are the Legendre polynomials. It is easy to show that

$$\frac{(2l+1)(2l'+1)}{2} C_{l\,1l} = l_> \text{ with } l_> = \begin{cases} l' \text{ for } l' > l \\ l \text{ for } l > l' \end{cases} \tag{77}$$

and we get for (76) the very simple expression

$$\sum_{L^T} (2L^T+1)[f_1(LlL'l'L^T)]^2 = \tfrac{1}{3} L_> l_> \tag{78}$$

and

$$\Omega_{ll'}^{\text{CB'I}}(nL, n'L') = \frac{16}{3} L_> [a_1[nL_1\, n'L']]^2 l_> \left| \int_0^\infty F_{kl} F_{k'l'} \frac{dr}{r^2} \right|^2. \tag{79}$$

The partial collision strength $\Omega_{ll'}$ is simply proportional to the atomic transition probability and to the probability that the incident electron, under the action of the Coulomb–field, goes from the kl state to the $k'l'$ state.

It is convenient to express $\Omega_{ll'}$ in terms of the Gaunt factors. In the same way that one defines the Gaunt factor by (6)

$$g(k, k') = \frac{8}{\sqrt{3}} \sum_{ll'mm'\mu} |(klm|r^{-2} Y_{1\mu}|k'l'm')|^2 \tag{80}$$

where

$$\phi(klm|\mathbf{r}) = Y_{lm}(\hat{\mathbf{r}})\, \frac{1}{r}\, F_{kl}(r)$$

and where F_{kl} is the Coulomb function (34), one can define a partial Gaunt factor, for each value of ll', by

$$g_{ll'}(k, k') = \frac{8}{\sqrt{3}} \sum_{mm'\mu} |(klm|r^{-2} Y_{1\mu}|k'l'm')|^2. \tag{81}$$

After angular integration,

$$g_{ll'}(kk') = \frac{2\sqrt{3}}{\pi} l_> \left| \int_0^\infty F_{kl} F_{k'l'} \frac{dr}{r^2} \right|^2 . \tag{82}$$

Writing the partial collision strength $\Omega_{ll'}$ and Ω_l in terms of $g_{ll'}$ and g_l, we obtain

$$\Omega_{ll'}^{\mathrm{CB'I}}(nL, n'L') = \frac{8\pi}{3\sqrt{3}} L_> [a_1[nL, n'L']]^2 g_{ll'}(k, k') \tag{83}$$

$$\Omega_l^{\mathrm{CB'I}}(nL, n'L') = \frac{8\pi}{3\sqrt{3}} L_> [a_1[nL, n'L']]^2 g_l(k, k') \tag{84}$$

with

$$g_l(k, k') = \sum_{l'} g_{ll'}(k, k').$$

In terms of the line strength

$$\mathcal{S}(nL, n'L') = 2L_> \left| \int_0^\infty P_{nL} P_{n'L'} r dr \right|^2 = 2L_> [a_1[nL, n'L']]^2 \tag{85}$$

we have

$$\Omega_l^{\mathrm{CB'I}}(nL, n'L') = \frac{4\pi}{3\sqrt{3}} \mathcal{S}(nL, n'L') g_l(kk') \tag{86}$$

and the total collision strength is

$$\Omega^{\mathrm{CB'I}}(nL, n'L') = \frac{4\pi}{3\sqrt{3}} \mathcal{S}(nL, n'L') g(kk'). \tag{87}$$

The Gaunt factor, which is of order unity, has been tabulated, e.g., by Grant (**6**). In the Bethe approximation Ω is simply proportional to the product of the line strength and the Gaunt factor $g(kk')$. If we use plane waves in (80) we have the well known expression

$$g(k, k') = \frac{\sqrt{3}}{\pi} \ln \left(\frac{k + k'}{k - k'} \right) \tag{88}$$

which gives the usual expression of the Bethe approximation for Ω. The Bethe approximation being frequently used, it is important to discuss in detail the limits of its validity.

We have to evaluate the respective importance of the two approximations (72) and (45). In Tables I and II we compare the results for the Bethe approximation and the results when we assume only the weak coupling (72)—approximation CB I.

For $k'^2 = 0$, we obtain

$$\text{for } 4s - 4p,\ k^2 - k'^2 = 0{\cdot}2309,\ \frac{\Omega(\mathrm{CB\,I})}{\Omega(\mathrm{CB'\,I})} = 0{\cdot}357$$

$$\text{for } 4p - 3d,\ k^2 - k'^2 = 0{\cdot}1062,\ \frac{\Omega(\mathrm{CB\,I})}{\Omega(\mathrm{CB'\,I})} = 0{\cdot}694.$$

This is easy to understand if we use the expression (86) for $\Omega_l^{\mathrm{CB'I}}$ in terms of the Gaunt factor $g_l(kk')$. The collision strength Ω_l varies like $g_l(kk')$, the variations of which are represented in figure 1 for the cases corresponding to the transition $4s - 4p$ and $4p - 3d$ with $k'^2 = 0$ and $k'^2 = 0{\cdot}1$. The Bethe approximation, and more precisely approximation (45), is especially good when the difference of energy $E_n - E_{n'}$ between the two atomic levels is small. When further calculations have been made for many other transitions, it should be possible to estimate the ratio $\frac{\Omega(\mathrm{CB\,I})}{\Omega(\mathrm{CB'\,I})}$ as a function of k' and $E_n - E_{n'}$. In such a way, it will be

TABLE I

Values of $\Omega_l(4s, 4p)$ *and* $\Omega(4s, 4p)$

l	$k'^2=0.0$						$k'^2=0.1$				
	CB I	CB II		CB′ I	CB′ II		CB I	CB II		CB′ I	CB′ II
		(2)	(3)		(2)	(3)		(2)	(3)		(2)
0	1·521	0·798	0·683	2·845	0·968	0·906	1·365	0·759	0·667	3·549	0·994
1	0·687	0·614	0·606	18·731	2·856	2·487	0·710	0·633	0·623	19·363	2·834
2	5·345	3·367	3·179	30·100	4·797	3·424	4·060	2·805	2·668	27·531	4·856
3	12·739	6·019	4·768	20·619	6·796	3·793	11·214	5·717	4·759	19·787	6·793
4	6·932	4·874	2·864	7·585	6·699	2·792	8·798	5·682	3·839	10·772	6·184
5	1·681	1·559	1·076	1·705	1·581	1·084	4·250	3·532	2·523	4·384	3·623
6	0·265	0·263	0·228	0·265	0·262	0·228	1·648	1·551	1·266	1·662	1·563
7	0·029	0·029	0·028	0·029	0·029	0·028	0·600	0·591	0·531	0·606	0·596
8	0·002	0·002	0·002	0·002	0·002	0·002	0·224	0·224	0·213	0·224	0·226
9							0·089	0·089	0·087	0·089	0·089
10							0·031	0·031	0·031	0·031	0·031
11							0·010	0·010	0·010	0·010	0·010
12							0·003	0·003	0·003	0·003	0·003
$\Omega=$	29·20	17·52	13·43	81·86	23·99	14·74	33·00	21·63	17·22	84·41	27·80

Note.—Columns headed (2) and (3) are approximations CB II and CB′ II calculated allowing for 2 levels ($4s$, $4p$) and for 3 levels ($4s$, $4p$ and $3d$).

TABLE II

Values of $\Omega_l(3d, 4p)$ *and* $\Omega(3d, 4p)$

	$k'^2=0.0$				$k'^2=0.1$		
	CB I	CB II	CB′ I	CB′ II	CB I	CB II	CB′ I
0	4·889	2·461	2·020	0·255	4·163	2·279	2·767
1	2·176	1·458	9·934	2·166	1·795	1·313	11·449
2	5·998	2·716	19·574	5·614	5·101	2·739	18·197
3	17·019	8·599	20·457	9·435	12·030	6·285	17·526
4	12·698	8·930	12·972	9·039	10·902	7·157	12·913
5	5·428	4·805	5·436	4·807	7·140	5·728	8·196
6	1·598	1·555	1·603	1·555	4·170	3·781	4·806
7	0·348	0·346	0·348	0·346	2·201	2·099	2·647
8	0·057	0·057	0·057	0·057	0·828	0·838	1·070
9	0·007	0·007	0·007	0·007	0·200	0·198	0·269
10					0·009	0·009	0·012
$\Omega=$	50·22	30·93	72·41	33·28	48·54	32·43	79·85

Note.—For the $3d-4p$ transitions in approximations CB II and CB′ II all three levels ($4s$, $3d$ and $4p$) are taken into account.

possible to correct empirically the cross sections quickly evaluated with the Bethe approximation CB′ I. Up to now we have assumed the weak approximation ; it remains to examine the case where we assume (45) but not (72).

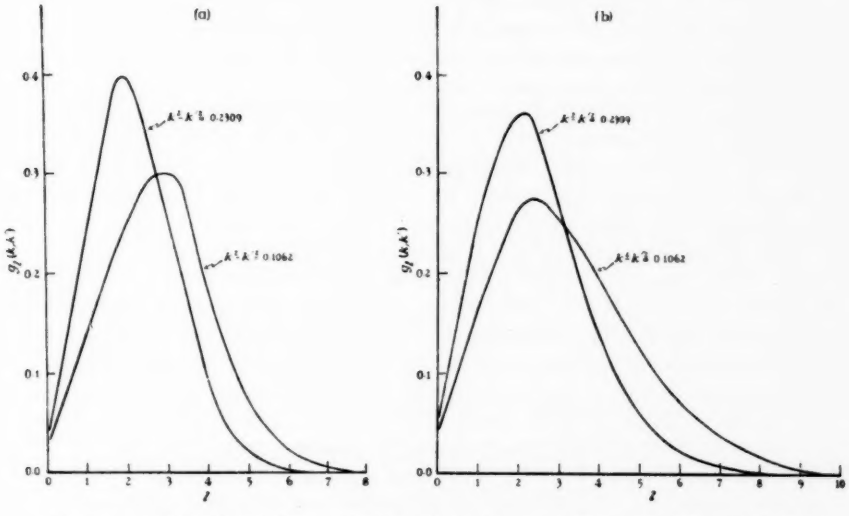


FIG. 1.—(*a*) $g_l(k, k')$ *for* $k'^2=0.0$, (*b*) $g_l(k, k')$ *for* $k'^2=0.1$. *The energy difference for* Ca^+ $4s-4p$ *corresponds to* $k^2-k'^2=0.2309$ *and that for* $3d-4p$ *to* $k^2-k'^2=0.1062$.

6. *The Bethe approximation for the* **R** *matrix.*—The elements of the **R** matrix are always given by (73) assuming (45), but for the transmission matrix

$$\mathbf{T}(\text{CB′ II}) = \frac{-2i\mathbf{R}(\text{Bethe})}{1-i\mathbf{R}(\text{Bethe})}. \qquad (90)$$

We can write (73) in terms of the Gaunt factor (82),

$$R_{aa'}(\text{Bethe}) = -\left(\frac{\pi}{\sqrt{3}}\right)^{1/2}[l_> L_>]^{1/2} f_1(LlL'l'\,;\,L^T)[\mathscr{S}(nL,n'L')g_{ll'}(k,k')]^{1/2}. \quad (91)$$

Thus the elements of the **R** matrix are easily calculated if we know $\mathscr{S}$ and if we have tables of f_1 and of the Gaunt factor.

T is computed as in the CB II method (Section 4), taking into account all the interactions between the $4s$, $4p$ and $3d$ levels. The results for Ω_l in the CB′ II approximation are included in Tables I and II. One sees that the results in the CB′ II approximation are in good agreement with those of the CB II approximation. This means that, for strong optically allowed transitions, it is possible to obtain good results using the CB′ II approximation, which is simple yet takes strong coupling effects into account.

At threshold one finds for the two transitions

$$\frac{\Omega(\text{CB′ II})}{\Omega(\text{CB′ I})} = \begin{cases} 0{\cdot}293 \text{ for } 4s-4p \\ 0{\cdot}460 \text{ for } 3d-4p. \end{cases} \quad (92)$$

For these two transitions the values of $\mathscr{S}$ are about the same and the differences in the ratios (92) for the two cases are therefore largely due to different values of the energy differences, $E_n - E_{n'}$. From Fig. 1 we see that for the $3d-4p$ transition there is a relatively larger contribution from the larger values of l, for which the CB′ I approximation is valid.

7. *Conclusions.*—From our study of the Ca^+ transitions we reach the following general conclusions:

(i) For optically allowed transitions in positive ions it is necessary to take many values of the angular momentum into account, up to $l \simeq 10$.

(ii) The Bethe approximation, with allowance for Coulomb distortion, has to be investigated from two points of view.

(*a*) From the weak coupling condition (72) which is valid or not according to whether the transition is weak or not. This depends on the line strength $\mathscr{S}$. (*b*) From the "long range" approximation (45). For l very small this is never valid and it is therefore necessary to consider the relative importance of the contribution from these values of l.

(iii) When the coupling is weak the approximations CB I and CB′ I may be used. The approximation CB′ I will overestimate the cross section by a factor of order two. This factor is smaller for small energy differences, larger for large energy differences. For transitions between excited states the effective atomic radii are greater and this disfavours the CB′ I approximation but since the energy differences will be small the CB′ I approximation may still give useful results.

(iv) When the coupling is strong the approximations CB II or CB′ II must be used. So long as allowance is made for strong coupling a relatively large contribution comes from large values of l and in consequence the CB II and CB′ II approximations give similar results. It therefore appears that the simple CB′ II approximation may give useful results for many strong optically allowed transitions.

Acknowledgment.—The author wishes to express his friendly gratitude to Dr M. J. Seaton, who suggested and supervised the work presented here.

Observatoire de Paris,
Meudon, Seine et Oise,
France:
1959 *December.*

Appendix I

The evaluation of the Coulomb integrals.

$$I_1 = \int_0^\infty F_{k_i l}(r) F_{k_f l'}(r) \frac{1}{r^2}\, dr.$$

For the case of an attractive field of two unit charges, the Coulomb functions F_{kl} are given by (34). It is convenient to compute first integrals of the type

$$M_l = \int_0^\infty F_{k_i l'}(r) F_{k_f l'}(r) \frac{1}{r}\, dr.$$

and to use the recurrence relations:

$$I_1(l \to l' = l+1) = \frac{k_f}{l+1}\,[(l+1)^2 + n_f^2]^{1/2} M_l - \frac{k_i}{l+1}\,[(l+1)^2 + n_i^2]^{1/2} M_{l+1} \qquad (1')$$

$$I_1(l \to l' = l-1) = \frac{k_i}{l}\,[l^2 + n_i^2] M_{l-1} - \frac{k_f}{l}\,[l^2 + n_f^2]^{1/2} M_l \qquad (2')$$

with $n_i = 1/k_i$ and $n_f = 1/k_f$.

From the expression for M_l given for the repulsive case by Alder *et al* (1) one finds, for the attractive case:

$$M_l = [k_i - k_f]^{-2}\, e^{+\frac{\pi}{2}[n_f - n_i]}\, \frac{(-x)^l}{\sqrt{n_i n_f}\,(2l+1)!}\, |\Gamma(l+1+in_i)||\Gamma(l+1+in_f)| \times \left(\frac{n_f - n_i}{n_f + n_i}\right)^{i(n_f+n_i)} F[l+1-in_i, l+1-in_f, 2l+2; x] \qquad (3')$$

with $x = -\dfrac{4n_i n_f}{(n_f - n_i)^2}$.

Using the properties of hypergeometric functions, one can express M_l as the real part of a function of $t = \dfrac{1}{1-x}$.

With

$$Q = \left(\frac{n_f - n_i}{n_f + n_i}\right)^{i(n_f+n_i)} F[l+1-in_i, l+1-in_f, 2l+2; x] \qquad (4')$$

we obtain

$$Q = 2\mathscr{R}e\, \frac{\Gamma(2l+2)\Gamma(i[n_i - n_f])}{\Gamma(l+1-in_f)\Gamma(l+1+in_i)}\, t^{l+1+i\frac{n_f-n_i}{2}} F[l+1-in_i, l+1+in_f, 1+in_f-in_i; t]. \qquad (5')$$

On the other hand, one can express Q in a convergent series in t of the form

$$Q = \sum_\lambda c_\lambda t^{\lambda+\kappa} \quad \text{with} \quad \kappa = l+1 \pm i\frac{n_f - n_i}{2}. \qquad (6')$$

After some transformations, one can write M_l in the form

$$M_l = \sqrt{2\pi}\, 2^{2l+1}[1 - e^{-2\pi(n_f - n_i)}]^{-1/2} (n_i n_f)^{l+1.5} (n_f - n_i)^{-1/2} (n_f + n_i)^{-(2l+2)} P_l(n_i n_f) \qquad (7')$$

with

$$P_l(n_i n_f) = x_0 + x_1 t + x_2 t^2 + \ldots$$

where

$$t = \frac{1}{1-x} = \left(\frac{n_f - n_i}{n_f + n_i}\right)^2. \qquad (8')$$

$P_l(n_i n_f)$ represents the real part of a sum of two complex conjugate series. One has

$$x_0 + iy_0 = e^{i(s+\epsilon_l)} \quad x_0 = \cos(s+\epsilon_l) \; y_0 = \sin(s+\epsilon_l) \tag{9'}$$

$$s = (n_f - n_i)[\log_e(n_f - n_i) - \log_e(n_f + n_i)] \tag{10'}$$

$$\epsilon_l = \arg\Gamma(l+1+in_f) - \arg\Gamma(l+1+in_i) - \arg\Gamma(1+i[n_f - n_i]) + \pi/2. \tag{11'}$$

The functions $\sigma_l = \arg\Gamma(l+1+in)$ satisfy the recurrence relation

$$\sigma_{l+1} = \sigma_l + \arctan\frac{n}{l+1} \tag{12'}$$

and σ_0 has been tabulated (**14**). If u_λ and v_λ are the real and the imaginary part of the recurrence coefficient ω_λ, with

$$\omega_\lambda = \frac{(\lambda+l+1)^2 + n_i n_f + i(n_f - n_i)(\lambda+l+1)}{(\lambda+1)^2 + i(\lambda+1)(n_f - n_i)} \tag{13'}$$

one has the recurrence relations

$$\begin{aligned} x_{\lambda+1} &= x_\lambda u_\lambda - y_\lambda v_\lambda \\ y_{\lambda+1} &= y_\lambda u_\lambda + x_\lambda v_\lambda . \end{aligned} \tag{14'}$$

Particular case: $k_f = 0$.

At the threshold, the evaluation of the M_l integral is considerably easier because of the confluence of the hypergeometric function F in (3'). One can show that, when $k_f \to 0$,

$$F\left[l+1 \mp in_i, l+1-in_f, 2l+2; -\frac{4n_i n_f}{(n_f - n_i)^2}\right] \to F[l+1-in_i, 2l+2, 4in_i]. \tag{15'}$$

From the study of the Coulomb functions of the form

$$P(x) = e^{-ix} F[l+1-in_i, \; 2l+2, \; 2ix] \tag{16'}$$

if we put

$$P(2n_i) = e^{-2in_i} F[l+1-in_i, \; 2l+2, \; 4in_i] \tag{17'}$$

it is possible to expand $P_l(2n_i)$ in the form

$$P_l(2n_i) = \sum_\lambda c_\lambda \frac{(2n_i)^\lambda}{\lambda!} \quad \text{with} \quad c_0 = 1 \;\; c_1 = \frac{n_i}{l+1}. \tag{18'}$$

The coefficients c_λ satisfy the recurrence relation

$$(2l+2+\lambda)c_{\lambda+1} = 2n_i c_\lambda - \lambda c_{\lambda-1}. \tag{19'}$$

After simplification of other factors in M_l, one finds

$$M_l = \frac{2\pi}{k_i^2} \frac{4^l}{(2l+1)!} n_i^{2l} e^{-\pi n_i} [A_{kl}]^{1/2} [1 - e^{-2\pi n_i}]^{-1/2} P_l(2n_i) \tag{20'}$$

with

$$A_{kl} = (1+k^2 l^2)(1+k^2(l-1)^2) \ldots (1+k^2) \quad \text{and} \quad A_{k0} = 1. \tag{21'}$$

Recurrence relations (1') and (2') become

$$I_1(l \to l' = l+1) = \frac{1}{l+1} M_l - \frac{k_i}{l+1} [(l+1)^2 + n_i^2]^{1/2} M_{l+1} \tag{22'}$$

$$I_1(l \to l' = l-1) = \frac{k_i}{l} [l^2 + n_i^2]^{1/2} M_{l-1} - \frac{1}{l} M_l. \tag{23'}$$

Appendix II

Numerical Calculations

W. Lawson and J. Lawson

It was required to obtain solutions of the equations

$$\left[\frac{d^2}{dr^2} - \frac{l(l+1)}{r^2} + \frac{2}{r} + k^2\right] F_{kl}(r) = 0 \qquad (1'')$$

and to evaluate the integrals

$$I_1 = \int_0^\infty F_{kl} F_{k'l'} \frac{1}{r^2}\, dr \qquad (2'')$$

and

$$I_2 = \int_0^\infty F_{kl} F_{k'l'} Z\, dr \qquad (3'')$$

where Z is a given function.

The calculations were carried out on the electronic computer of the Department of Physics, University College, London. All of the integrals I_1 were evaluated using the analytic expressions obtained in Appendix I and, as a check, some of these integrals were also calculated by numerical integration.

Numerov's method was used for integrating the equation (1''). Putting $F'' = uF$, $F_n = F(hn)$ and $u_n = u(hn)$ where h is the interval, we have

$$F_{n+1} = \frac{A_n F_n - B_{n-1} F_{n-1}}{B_{n-1}} \qquad (4'')$$

where

$$A_n = \left(2 + \frac{10h^2}{12} u_n\right) \quad \text{and} \quad B_n = \left(1 - \frac{h^2}{12} u_n\right). \qquad (5'')$$

Starting solutions were obtained from expansions of the form

$$F_{kl} = C_{kl} r^{l+1} [1 + a_1 r + a_2 r^2 \ldots] \qquad (6'')$$

where (compare (35))

$$C_{kl} = \left(\frac{\pi}{2}\right)^{1/2} \frac{2^{l+1}}{(2l+1)!} \left[\frac{A_{kl}}{[1 - e^{-2\pi/k}]}\right]^{1/2}. \qquad (7'')$$

The integrations were carried out with an interval $h = 0{\cdot}1$ for $0 \leqslant r \leqslant 20{\cdot}2$, and $h = 0{\cdot}2$ for $r > 20{\cdot}2$. A substantial region of overlap provided a check on the integration routine. Simpson's rule was used to evaluate the integrals I_1 and I_2.

The programme for solving equation (1'') was stored in duplicate so as to provide the required pair of functions F_{kl} and $F_{k'l'}$ in one cycle. An arithmetic check was included in the programme which ensured that the difference equation

$$\delta^2 F = h^2 (F'' + (1/12)\delta^2 F'')$$

was satisfied to a given degree of accuracy at the end of each cycle. If the check was not satisfied the cycle was repeated, otherwise the values obtained for F_{kl} and $F_{k'l'}$ were used as input to a further programme which formed and accumulated the products

$$F_{kl}F_{k'l'}(1/r^2)S \text{ and } F_{kl}F_{k'l'}ZS \left(S=\frac{4h}{3} \text{ or } \frac{2h}{3}, \text{ the Simpson factor}\right).$$

The functions and part integrals could be printed out at any value of the argument r.

This method of building up the integrals I_1 and I_2 as the solutions of (1″) proceed has the following advantages:

1. It is economical in storage space.

2. The last few consecutive values of the part integral provide a useful check on the degree of convergence.

3. The entire calculation may be easily extended if convergence is not satisfactory.

An alternative, which would probably be preferred on larger and faster machines, would be to form and store all required functions before proceeding to evaluate I_1 and I_2 by some sub-routine.

The accuracy of the calculations was checked in various ways. For $k=0$ the solutions of (1″) are known to be

$$F_{0l}(r)=\sqrt{\pi r}J_{2l+1}(\sqrt{8r}). \tag{8″}$$

Comparisons were made using tables of Bessel functions at several convenient values of r. The order of accuracy of F_{0l} was 5 s.f. for $l=0$ decreasing to 3 s.f. for $l=6$. For $k>0$ the asymptotic form (34) is a close approximation only in the limit of very large r. The calculated functions were therefore compared, for moderately large r, with WBK functions having correct asymptotic form. These are

$$F_{kl}(WBK)=g^{-1/2}\sin\left[\int^r g(r')\,dr'\right]$$

where

$$g(r)=\left[\frac{2}{r}-\frac{l(l+1)}{r}+k^2\right]^{1/2}. \tag{9″}$$

The comparison showed that, for $r\sim 60$, the accuracy of F_{kl} for $k>0$ was comparable with that of F_{0l}. These checks proved to be invaluable in detecting mistakes in the hand calculated starting solutions. A mistake in the multiplier C_{kl} could easily be corrected without repeating the whole calculation.

The general falling off of accuracy with increasing l is almost certainly due to the fact that the starting values for large l were small in magnitude and therefore contained few significant figures. The larger the value of r at which the solution is started the more terms are needed in the expansion, and a compromise was sought which would produce the final accuracy desired. Bearing in mind that the integrals I_1 and I_2 decrease in magnitude as l increases, it was possible to tolerate a fall in accuracy of this order.

In the case of the integrals I_1 which was required for checking purposes, the accuracy of the numerical integrations was often limited by the degree of convergence obtained.

References

1. Alder, K., Bohr, A., Huus, T., Motteson, B., and Winther, A., *Rev. Mod. Phys.* **28,** 432 (1956).
2. Biedenharn, L. C. *Tabulation of the radial coulomb integral.* AEC Report LA 2106.
3. Biedenharn, L. C., McHale, J. L., and Thaler, R. M., *Phys. Rev.* **100,** 376 (1955).
4. Burgess, A. (to be published).
5. Condon, E. V., and Shortley. *The Theory of Atomic Spectra.* Cambridge University Press.
6. Grant, I. P., *M.N.* **118,** 241 (1958).
7. Hartree, D. R., and Hartree, W., *Proc. Roy. Soc.* A, **164,** 167 (1938).
8. Jefferies, J. T., *Austr. J. Phys.*, **7,** 22 (1954).
9. Mott, N. F., and Massey, H. S. W. *The Theory of Atomic Collisions*, Oxford, 1949.
10. Percival, I. C., and Seaton, M. J. *Proc. Cambridge Phil. Soc.*, **53,** 654 (1957).
11. Seaton, M. J., *Proc. Phys. Soc.*, **68,** 457 (1955).
12. Seaton, M. J. *Proc. Phys. Soc.* (to be published).
13. Seaton, M. J. *Rev. Mod. Phys.*, **30,** 979 (1958).
14. Tables of Coulomb Wave Functions, National Bureau of Standards, Applied Math. Series 17.

PHOTOMETRY IN THE MAGELLANIC CLOUDS, III. THE CLUSTER NGC 1783

Allan R. Sandage and Olin J. Eggen

(Received 1960 January 27)

(Communicated by the Astronomer Royal)

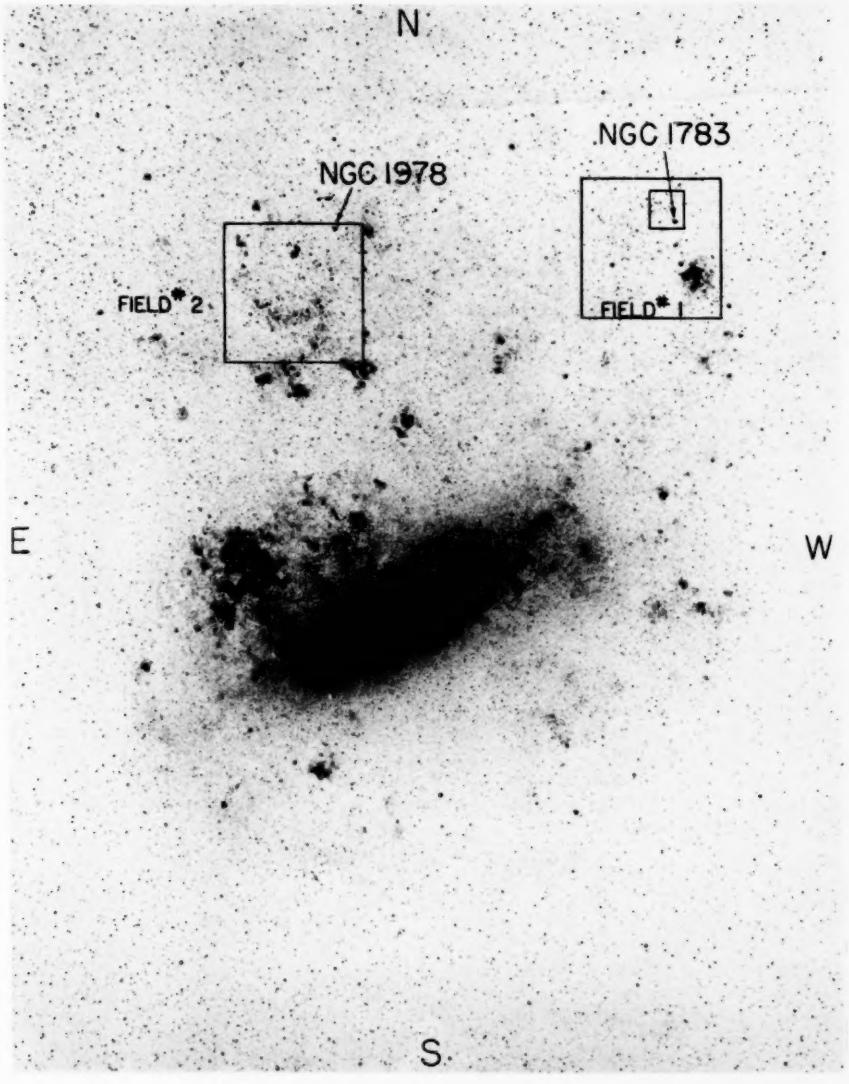
Summary

Photometry is reported of the cluster NGC 1783 in the Large Magellanic Cloud. The 74-inch Radcliffe reflector at Pretoria, South Africa, was used. NGC 1783 had previously been suspected as being a " normal " globular cluster similar to the halo clusters in our Galaxy but the $(V, B-V)$-diagram shows that this is not the case. The red giant sequence has a much more gradual slope than that of halo clusters, and the giant sequence extends to an extremely red terminus point at $(B-V)_0 = +2{\cdot}4$. The diagram, although dissimilar to globulars in our Galaxy, is almost identical with the $(V, B-V)$-diagram of NGC 361 and NGC 419 obtained by Arp in the SMC and to the diagram of NGC 1978 obtained by Hodge in the LMC.

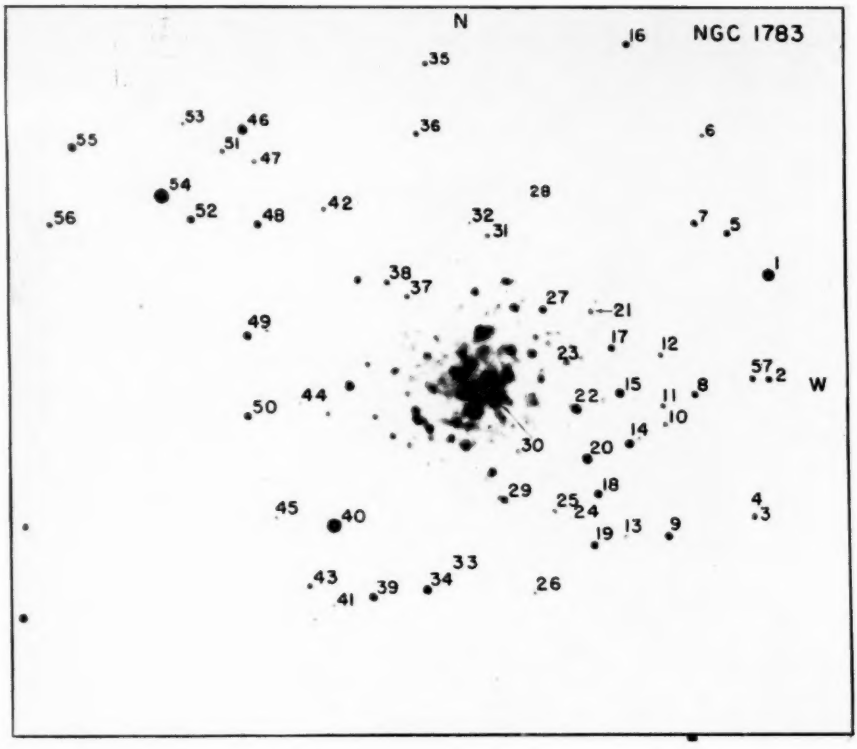
A composite diagram is shown which compares the $(V, B-V)$-diagram of M3, M92, NGC 6356, and NGC 1783. The position and slope of the giant branch of 1783 suggests that its cluster stars may be very rich in heavy elements, similar to NGC 6356. But this appears to be contradicted by the extreme redness of the giant branch. Spectrograms of the cluster or of its individual stars would appear to be necessary to settle the question. If clusters similar to the known clusters in our Galaxy are not found in the Magellanic Clouds, then the quest for the cosmic distance scale is made more difficult because of the inapplicability of the principle of uniformity to the stellar contents of galaxies.

Introduction.—NGC 1783 is a rich, compact cluster in the Large Magellanic Cloud (LMC), located about $4\frac{1}{2}$ degrees north-west of the centre of the bar. An early catalogue of Shapley (1930) lists NGC 1783 as a globular cluster. More recently, Gascoigne and Kron (1952) measured the colour index and the integrated magnitude and showed that the cluster was red. Because of (1) the appearance on photographic plates, (2) the fact that the brightest stars are red, (3) Gascoigne and Kron's value of the integrated colour index of $P-V=0{\cdot}48$ $(B-V\approx 0{\cdot}60)$, and (4) their value for the absolute magnitude of about $M_p \approx -8$, NGC 1783 was suspected to be a normal globular cluster like the halo clusters in our Galaxy. Consequently, this cluster was placed on the current programme of photometry (Eggen and Sandage 1960; Woolley 1960) to provide data for the calibration of the distance to the clouds.

The location of NGC 1783 is shown in Plate 3, which is from an Hα photograph taken by Henize with the Mt Wilson 10-inch refractor. This telescope was taken to South Africa in 1949. The boxed area of Plate 3 outlines the region of the photoelectric sequence in LMC field I which is identified in Paper I of this series (Eggen and Sandage 1960, Plate 1). NGC 1783 is situated only 6′ south of the



Identification of the region of NGC 1783. The boxed areas are LMC fields 1 and 2 shown in Plates 1, 2 and 3 of Paper I. The reproduction is from an Hα photograph taken by K. Henize with the Mt Wilson 10-inch refractor with an exposure of 315 minutes.



Identification chart for the programme stars in NGC 1783. The magnitudes and colours of the stars are given in Table I of the text.

standard photoelectric sequence of LMC field I. Six minutes of arc is small enough so that both the cluster and the sequence stars appear within the coma-free field of plates taken at the Newtonian focus of the 74-inch Radcliffe reflector. When used for photography, the 74-inch mirror was usually diaphragmed to 43 inches to increase the photometric field.

The photometry.—Many plates of NGC 1783 were taken during the 1958–59 observing season. Due to the intense background field of the LMC itself, only the shorter exposure plates were selected for measurement. Two blue plates (Eastman 103*a*O + 2 mm of Schott GG 13) and two yellow plates (Eastman 103*a*D + 2 mm of Schott GG 11) were measured with the Sartorius iris-diaphragm photometer at the Royal Greenwich Observatory. The blue plates were both of 20 minute exposure, the two yellow plates were 7 and 15 minute exposure. 58 stars in the neighbourhood of NGC 1783 were measured. These are identified in Plate 4. The stars were chosen as close to the cluster centre as possible without encountering serious crowding difficulties.

TABLE I

Colours and magnitudes for stars in and near NGC 1783

Star	V	$B-V$	$AD(V)$	$AD(B)$	Star	V	$B-V$	$AD(V)$	$AD(B)$
	m	m	m	m		m	m	m	m
1	14·90	0·80	0·00	0·02	31	16·68	1·27	...	...
2	15·96	1·77	...	0·02	32	16·74	1·22	...	...
3	16·41	1·46	...	0·02	33	17·01	1·02	...	...
4	16·76	0·93	...	0·06	34	15·59	2·14	...	0·10
5	15·98	2·01	0·03	0·09	35	16·35	1·56	...	0·02
6	16·83	1·20	...	0·03	36	16·16	1·74	...	0·05
7	15·98	0·13	0·04	0·04	37	16·34	1·36	...	0·00
8	16·03	1·82	...	0·03	38	16·44	1·38	...	0·02
9	15·84	1·90	0·04	0·05	39	15·71	1·95	0·05	0·08
10	16·57	1·35	...	0·06	40	14·50	0·73	0·03	0·07
11	16·46	1·52	...	0·00	41	16·70	1·24	...	0·01
12	15·51	2·34	...	0·02	42	16·68	−0·18	...	0·05
13	16·83	1·18	...	0·01	43	16·56	1·26	...	0·08
14	15·69	2·06	0·09	0·14	44	16·96	0·94	...	...
15	15·71	2·02	0·01	0·02	45	16·78	1·12	...	0·00
16	15·94	1·16	0·02	0·01	46	15·41	0·79	0·04	0·06
17	16·31	1·48	...	0·01	47	16·62	1·24	...	0·02
18	15·80	1·93	0·03	0·08	48	15·72	0·92	0·01	0·04
19	16·03	1·65	...	0·06	49	15·71	1·95	0·07	0·08
20	15·49	0·72	0·07	0·08	50	15·84	1·35	0·04	0·04
21	16·82	−0·03	...	0·03	51	16·21	1·76	...	0·00
22	15·60	1·96	0·08	0·02	52	15·70	1·95	0·01	0·00
23	16·44	1·39	...	...	53	16·78	1·21	...	0·01
24	17·05	0·94	...	0·01	54	14·12	0·95	0·01	0·00
25	16·82	0·92	...	0·03	55	15·60	1·99	0·02	0·05
26	16·96	1·09	...	...	56	16·20	1·59	0·17	0·01
27	15·98	1·77	0·02	0·04	57	16·07	0·83	...	...
28	17·09	0·91	...	...	58	15·68	1·85	0·02	0·00
29	16·36	1·34	...	0·04					
30	14·85:	−0·05:	...	...					

The magnitudes of the 58 unknown stars were read from calibration curves made with the photoelectric standards. The resulting V and $B-V$ values are given in Table I together with the deviations from the mean for those stars which were measured twice. These deviations suggest that the tabulated values of Table I have probable errors of $\epsilon_V = \pm 0.024$ mag., $\epsilon_B = \pm 0.022$ mag. and $\epsilon_{B-V} = \pm 0.032$ mag. No systematic error is expected in the photometry because the standard stars, read back through the calibration curves, gave magnitudes which differed by $V_{pe} - V_{pg} = -0.004$ mag. and $B_{pe} - B_{pg} = +0.022$ mag. from the photoelectric values, over the entire range. The photographically smoothed values for the standards are given in Table II where a comparison with the photoelectric data can be made. The photographic values in Table II are in excellent agreement with the data of Table V in Paper I which were determined from an entirely different set of plates.

TABLE II

Values of the Photoelectric Standards in Field I determined photographically

Star*	V_{pe}	V_{pg}	B_{pe}	B_{pg}
a	9·64	9·64	10·66	10·66
34	11·14	11·12	11·41	11·46
23	12·08	12·13	12·63	12·63
38	12·17	12·12	12·06	11·95
35	12·68	12·64	14·59	14·52
7	13·12	13·06	13·95	13·86
33	13·42	13·57	13·24	13·30
36	13·50	13·18	13·23	13·05
1	13·90	13·89	13·84	13·82
27	14·50	14·51	14·20	14·21
26	15·15	15·32	14·95	15·05
11	15·16	15·21	14·99	14·91
5	15·30	15·14	15·83	15·85
25	15·54	15·71	15·45	15·59
28	15·73	15·69	15·59	15·61
31	15·80	15·80	15·68	15·72
6	15·90	15·75	15·87	15·59
24	16·62	16·61	17·87	17·81
37	17·07	17·07	17·04	17·08

* These stars are identified in Plate 1 of Paper I, *M.N.*, **120**, 79, 1960.

The $(V, B-V)$*-diagram.*—The colour-magnitude diagram is plotted in Fig. 1. A well defined red giant sequence is present, together with 13 stars which scatter from this sequence. These 13 stars undoubtedly belong to the field and are not members of the cluster itself. Unfortunately, since our photometry extends to only $V = 17{\cdot}07$, $B = 17{\cdot}87$,we cover only the top $1\frac{1}{2}$ magnitudes of the diagram for NGC 1783. However, several unique features are present which indicate that NGC 1783 is unlike any normal globular cluster known in our Galaxy. These features are (1) the giant branch extends to $B-V = 2{\cdot}4$ and is well populated to its terminus. In all globular clusters studied so far in our own Galaxy, the giant branches end abruptly at $B-V \approx +1{\cdot}6$. Very few stars as red as $(B-V)_0 = +2{\cdot}4$ are known in the Galaxy and none at all are known in the halo globular clusters. (2) The slope of the giant sequence of NGC 1783 is much less steep than the giant branches of normal clusters. The gradient of the sequence is $\Delta V/\Delta(B-V) = 1{\cdot}25$. Over the same range of colour indices ($B-V = +1{\cdot}6$ to

$B-V=0{\cdot}8$) the average gradient of M3 is 2·9 although the giant branch in M3 cannot be represented by a straight line as can NGC 1783. (3) There is no drop towards a vertical subgiant sequence beginning at $(B-V)_0=+0{\cdot}9$ as in M3 or M92. To the limit of our data, the giant sequence continues in a linear fashion.

These results are not entirely new. Arp has found similar features in the two supposed globular clusters NGC 361 and NGC 419 in the SMC (Arp 1958b, 1958c). In fact, Arp's diagram for NGC 419, plotted from Table III of his paper (1958b), is nearly identical with Fig. 1 of the present paper if we correct for the +0·08 mag. difference in the zero point of the V magnitudes between our photometric transfer and Arp's transfer from the northern standards to the Clouds (Eggen and Sandage 1960). The giant sequence of NGC 1783 reaches to a slightly redder colour-index and to a somewhat brighter visual magnitude than that of NGC 419 but Arp states (private communication) that, if the brightest stars near the centre of NGC 419 were included in his diagram, this giant sequence would mimic that of NGC 1783, reaching $B-V\approx+2{\cdot}4$ at $V\approx15{\cdot}5$.

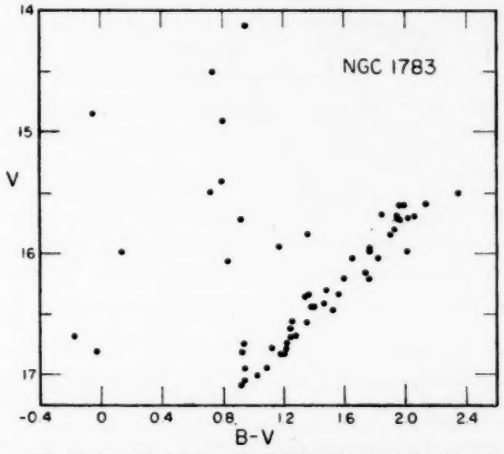


FIG. 1.—*The* $(V, B-V)$*-diagram for NGC* 1783 *from the data of Table I.*

Similarly, P. W. Hodge (1960) has just completed a study of the $(V, B-V)$-diagram of NGC 1978 in the LMC using ADH Schmidt plates and a sequence stepped down from Arp's photoelectric sequence near NGC 1866 (Arp, Sandage, and Thackeray, in preparation). His diagram is very similar to Fig. 1 of this paper and is unlike globular clusters in our Galaxy. The slope of the giant branch in NGC 1978 is definitely less steep than that of M3.

Arp carried the diagram of NGC 419 to very faint magnitudes: $V\approx19{\cdot}6$, $B\approx20{\cdot}1$. A vertical branch was found, beginning at $V\approx17{\cdot}6$, which is similar to the subgiant region of M3 or M92. This is just beyond the limit of our photometry in NGC 1783. In addition to the above differences in the giant branch, Arp found that NGC 419 had a gap of 0·4 mag. in the subgiant branch extending from $V=17{\cdot}2$ to $V=17{\cdot}6$. Further, he found that NGC 419 had no horizontal branch. But with these exceptions, the diagram of NGC 419 was generally similar to clusters in our Galaxy and Arp concluded that NGC 419 was " globular-cluster like ".

These studies indicate that *none* of the supposed globular clusters yet studied in detail in either Cloud (NGC 361, 419, 1783, and NGC 1978) can be identified in all their peculiar properties with clusters in our own Galaxy. Consequently, use of these clusters as distance indicators would seem to be premature at the moment.

In an effort further to compare NGC 1783 with clusters in the Galaxy, we have superposed the giant branch shown in Fig. 1 with a composite diagram of several representative globular clusters. Although such a comparison is uncertain at present because we probably do not yet know the correct modulus to the LMC,

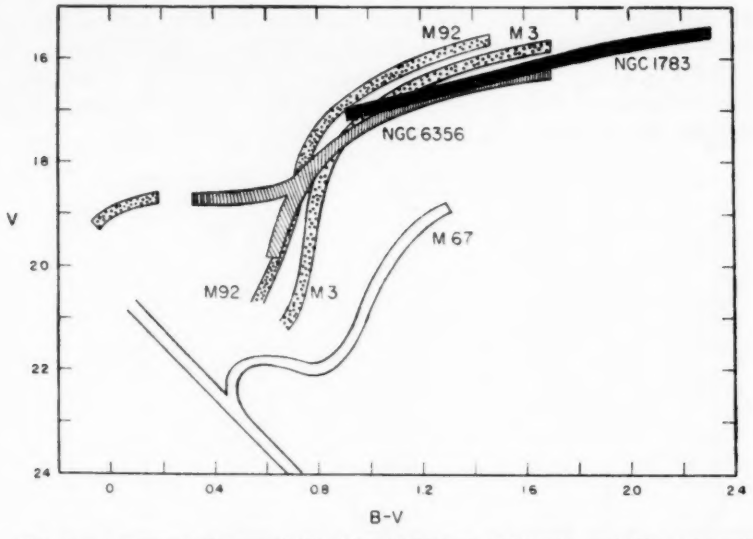


FIG. 2.—*A composite diagram of M3, M92, NGC 6356, NGC 1783, and M67. The ordinate is apparent visual magnitude as seen at the distance of the LMC. The modulus of the LMC is assumed to be* $m-M=18{\cdot}7$*; the absolute magnitudes of the horizontal branches of the clusters are assumed to be* $M_V=0{\cdot}00$*. The relative position of NGC 1783 with respect to the other clusters will change if either the value for* $m-M$ *or of* M_V *is changed.*

the resulting diagram defines the limits of our knowledge. Fig. 2 combines the $(V, B-V)$-diagram for M92 (Arp 1955; Sandage and Walker 1956, unpublished), M3 (Sandage 1953; Johnson and Sandage 1956), NGC 6356 (Sandage and Wallerstein 1960), and NGC 1783. This figure is intended only to show the relative shapes of the giant sequences and is based on several assumptions, some of which are likely to be incorrect.

1. The horizontal branches of M3, M92 and NGC 6356 are all assumed to be at $M_V=0^m{\cdot}0$. However the horizontal branch of the $(M_V, B-V)$-diagram for stars in M3 is more likely nearer $M_V\approx+0^m{\cdot}5$ (Eggen and Sandage 1959) and the true positions of those for M92 and NGC 6356 are at present unknown.

2. The true modulus of the LMC is taken as $m-M_V=18^m{\cdot}7$. This modulus is based on the combination of (*a*) $B=17^m{\cdot}8-2{\cdot}2\log P$ for the apparent (median) magnitude-period relation for cepheids in the SMC (Arp 1958a) and (*b*) $M_B=-0^m{\cdot}9-2{\cdot}2\log P$ for the preliminary period-luminosity relation of the cepheids in galactic clusters (Arp, Kraft and Sandage, in preparation). The

two Clouds are assumed to be at the same distance and, perhaps more questionably, the galactic and Cloud cepheids are assumed to have the same period-luminosity relation.

3. NGC 1783 is assumed to be unreddened. Although the extreme redness of stars in this cluster might be explained as due to a reddening of $E_{B-V} \approx +0^{m}.6$, this seems unlikely because of the presence in the cluster region of stars near the theoretical blue limit of $B-V=-0^{m}.4$; No. 42 in Table I with $B-V=-0^{m}.18$ and several stars immediately north of the cluster with $B-V=-0^{m}.4$ (Woolley 1960).

If we adopt the identification of NGC 1783 with NGC 6356 on the basis of the agreement of the slopes of their giant branches, then we might expect the members of NGC 1783 to have the same high metal abundance that is indicated by the spectra of stars in NGC 6356 (Sandage and Wallerstein 1960). However, if this were so, the question arises as to why the colour indices of the stars in NGC 1783 are not limited to $B-V<+1^{m}.7$ by the strong TiO bands which should be present if the abundance of metals is high. It is obvious that (1) integrated spectra of the entire cluster, or (2) spectra of the individual members of NGC 1783 or (3) measurement of the ultra-violet excess of cluster members would be of the greatest importance in answering this question.

Perhaps the only conclusion to be drawn from the present data is that no cluster in our own Galaxy is known to be similar to NGC 1783 in all its features. Because NGC 1783 is similar to three other clusters in the Magellanic Clouds (NGC 361, 419 and 1978) it would seem important to search for "normal" clusters in the Clouds and for objects that might match NGC 1783 in our own Galaxy. If no such systems are found, then the similarity of the stellar contents of galaxies must be seriously questioned and the search for the cosmic distance scale will be more difficult than once imagined.

Mount Wilson and Palomar Observatories,
Carnegie Institution of Washington,
California Institute of Technology,
Pasadena, California:
1960 *January* 1.

Royal Greenwich Observatory,
Herstmonceux Castle,
Sussex.

References

Arp, H. C., 1955, *A.J.*, **60**, 317.
—— 1958a, *A.J.*, **63**, 45.
—— 1958b, *A.J.*, **63**, 273.
—— 1958c, *A.J.*, **63**, 487.
Arp, H. C., Sandage, A. R., and Thackeray, A. D., 1960, in preparation.
Eggen, O. J., and Sandage, A. R., 1959, *M.N.*, **119**, 255.
—— 1960, *M.N.*, **120**, 79 (Paper I).
Gascoigne, S. C. B., and Kron, G. E., 1952, *Pub. A.S.P.*, **64**, 196.
Hodge, P. W., 1960, *Ap. J.*, **131**, 351.
Johnson, H. L., and Sandage, A. R., 1956, *Ap. J.*, **124**, 379.
Sandage, A. R., 1953, *A.J.*, **58**, 61.
Sandage, A. R., and Wallerstein, G., 1960, *Ap. J.*, **131**, 598.
Shapley, H., 1930, *Star Clusters*, Table XIII, p. 187, McGraw-Hill Book Company, New York.
Woolley, R. v. d. R., 1960, *M.N.*, **120**, 214 (Paper II).

THE FORMATION OF POPULATION I STARS, PART II. THE FORMATION OF MOLECULAR HYDROGEN IN INTERSTELLAR MATTER

W. H. McCrea and D. McNally

(Received 1960 March 28)

Summary

This paper is essentially an investigation of the possibility of molecule-formation (chiefly H_2) by surface-reactions on interstellar dust-grains. The process has often been mentioned but here we seek to derive as precise an expression as possible for the rate of formation in terms of the properties of the grains. The effect upon the rate of a collision between interstellar clouds is also considered. The case of interstellar CH is discussed; it seems that some of the difficulties encountered by Bates and Spitzer may be overcome if we recognize the possibility of the enhanced production of CH in the conditions produced by a collision between clouds. If this is accepted, it indicates that the type of process here considered does actually take place with a fairly high efficiency. The case of H_2 is then considered. An efficiency comparable to that inferred in the case of CH would give a proportion of H_2 in ordinary interstellar matter in satisfactory agreement with Kahn's estimate of the amount required to cool the matter after heating by cloud-collisions. Such an efficiency would also explain why free atomic hydrogen would not be found with a density more than the order 100 atom cm^{-3}. If H_2 is produced at the inferred rate, then if an ordinary interstellar cloud is compressed by a factor about 100 the material will mostly take the form of H_2. This is of much significance for the process of star formation. Moreover, this significance is likely to remain, even if the process of H_2-formation has about the lowest efficiency that could be admitted.

---

1. *Introduction.*—In this paper we are concerned with interstellar matter only in H I regions. Such matter almost certainly consists predominantly of hydrogen gas. It is usually assumed that the hydrogen is mainly in (neutral) atomic form rather than in molecular form. At present this assumption cannot be tested directly by observation because molecular hydrogen provides no spectral features that would permit its detection from beneath the Earth's atmosphere. However, the assumption can apparently be justified by the generally self-consistent account of the properties of the material to which it leads. This account would be invalidated were a large proportion of the hydrogen in molecular form. For instance, in that case the material would be at a lower kinetic temperature than that actually inferred in several ways. Also, were molecular hydrogen abundant, then we should expect other molecular species, that can be observed, to be much more abundant than they are found to be. Kahn has indeed discussed the problem and has estimated that roughly about 0·5 per cent of interstellar

hydrogen may be in molecular form. Nevertheless none of these arguments excludes the possibility that molecular hydrogen is relatively plentiful in certain parts of the interstellar medium.

The significant consideration is that with respect to molecule-formation the hydrogen in H I regions has to be regarded as *super-cooled*. This is in the sense that, were it to attain thermodynamic equilibrium at anything like its kinetic temperature, it would be almost entirely in molecular form. In other words, were the hydrogen to be all converted to molecular form, it would remain in this form except to the extent to which photo-dissociation by starlight would operate to reduce it.

We must therefore conclude in the first place that the hydrogen in H I regions "originates" in atomic form. Then the reason why it remains largely in this form, at any rate in large regions, is that all processes of combination into molecular form are exceedingly slow under the conditions obtaining in these regions.

The rates of all processes of combination must, however, be proportional to at least the second power of the density. Now it is generally considered that Population I stars are formed from raw material that is at some stage the interstellar matter of H I regions. This demands an increase in density of the material concerned by an enormous factor of the order 10^{23} at least. The question arises as to whether any mechanism of combination can become so effective at some phase of the condensation process that we have to contemplate the possibility of all the hydrogen becoming molecular at that stage. Clearly such a possibility could have a profound effect upon the subsequent course of the condensation process.

The main purpose of the present paper is to discuss the problem from the standpoint of the possibility just mentioned. In the past, the problem has been mentioned mostly from the standpoint of the possible existence of molecular hydrogen in more or less steady conditions rather than that of its direct involvement in processes of evolution.

It would naturally be desirable if possible to estimate both upper and lower bounds to the rate of molecule-formation. Unfortunately not enough is as yet known quantitatively about the possible processes for useful results to be got in this way. Consequently we are compelled to adopt a somewhat different approach. After briefly reviewing the several possible processes, we consider in some detail one particular process that would apparently be the most effective were conditions "favourable" in a sense to be stated. We then examine the consequences that would follow if these conditions are actually realized. We attempt *a posteriori* to say whether these consequences indicate that the conditions probably are realized and also to say whether very different consequences would have to be expected if they are not.

2. *Possible processes for producing* H_2.—The simplest process for producing a hydrogen molecule is an encounter between two hydrogen atoms with the emission of a photon. Although the relevant collision cross-sections are not known, they are believed to be so small that the process would be ineffective in normal H I regions (Herzberg 1955). Also, as will be referred to again in Section 3, the corresponding processes are estimated to be quite inadequate to account for the observed numbers in the case of other molecules.

Herzberg (1955) has also suggested that certain chemical exchange processes such as $CH + H \rightarrow C + H_2$ might be operative. Again the cross-sections are not known but they would have to be surprisingly large if such processes play a significant role in H_2-production.

The only type of encounter essentially different from those just mentioned in which a hydrogen atom collides with an atom or a molecule is that in which it collides with a dust-grain. In fact it is generally considered that the most likely process of molecule-formation is a surface reaction on the grains of interstellar dust as was first suggested by van de Hulst (1948). For it is known that if any solid is exposed to an atomic gas some of the atoms adhere to its surface. Also it is known that further atoms striking the surface can combine with those already adhering, the energy released in the process being taken up in the first instance by the solid rather than being directly emitted as a photon. It is such a process that we consider below; but mention must first be made of molecules other than those of hydrogen.

3. *Other molecules.*—Bates (1951) has computed rate-coefficients for the formation by radiative association of certain molecules including CH, CH^+. Using these values, Bates and Spitzer (1951) investigated the equilibrium densities of CH and CH^+ in interstellar space. They concluded that the rates of association are too small by a factor of 500 to 1000 to account for the observed densities.

They then suggested that interstellar CH molecules may be formed from CH_4 evaporating from dust grains in the vicinity of hot stars. However, they did not suggest that such a process could account for the presence of CH molecules in ordinary H I regions.

4. *Formation of H_2 on grains.*—We accept it as a fact that a solid grain must at any instant have some atoms of interstellar hydrogen adhering to its surface. Chemists recognize the two possibilities of such atoms adhering at particular sites and of their having some mobility on the surface. This difference need not concern us. For it would only affect the way in which the probability ϖ defined below might be further analysed, were this necessary for other purposes.

It appears that the energy required to remove an adhering atom is generally of the order of 1 eV. Since the surrounding gas has a kinetic temperature about 100 degrees, an atom of the gas has energy only of the order of 10^{-2} eV. Consequently, there is no question of an adhering atom being merely knocked off the surface by atomic collisions.

Thus, when the grain is in any given condition, a further hydrogen atom impinging on its surface may only either rebound (or change places with an adhering atom which is the same thing for present purposes), or adhere to the surface, or combine with an already adhering atom to form a H_2 molecule.

If the impinging atom combines with a surface atom the process releases 4·47 eV which, as would be expected, is more than ample to detach the resulting molecule from the grain. We shall in fact assume that the molecule is certain sooner or later to become detached. For there cannot be progressive building-up of molecular hydrogen on the grain unless its temperature is below the melting-point of solid molecular hydrogen. This is 14 deg., while the grains are estimated to have a temperature of 20 deg. (Allen 1955). But in any case (as has been

pointed out to us by Dr F. D. Kahn) at the low gas-density concerned, a grain of solid molecular hydrogen would probably sublimate rather than grow. If then the molecules do not leave the surface, the only remaining alternative would be to suppose the process of molecule-formation to be halted by the surface atoms becoming replaced by surface molecules. This is not to be expected since the affinity of the surface for atoms is bound to be greater than its affinity for molecules.

These general considerations appear to be in good agreement with the theoretical and experimental studies of Smith (1943) who has given a general discussion of this type of surface reaction. Incidentally, it is easy to verify from his experimental results that the molecules formed on a surface do become detached. For it can be estimated that in the course of a single experiment each of the "sites" on the catalytic surface used as a probe must have been re-occupied about 10^{10} times over per second.

We define ϖ to be the mean probability that hydrogen atom impinging on a grain will in consequence form part of a H_2 molecule. From what has just been said $\frac{1}{2}\varpi$ is then also the probability that a collision between an atom and a grain results in the addition of a H_2 molecule to the surrounding gas.

The greatest possible value of ϖ is unity. In the most favourable case half the collisions will produce molecules and the other half will serve to replenish the surface-layer.

5. *Rate of formation of H_2.*—For the hydrogen gas we write,

H = atomic mass.

n_1 = number of free atoms per unit volume.

n_2 = number of hydrogen molecules per unit volume.

$n_0 = n_1 + 2n_2$ = total number of free and combined atoms per unit volume.

T = kinetic temperature, assuming a Maxwellian velocity-distribution.

We also define μ by

$\mu n_0 H$ = total mass of interstellar material per unit volume.

For the grains we write

N = number of grains per unit volume.

r_1 = mean radius of grain.

$A = 4\pi N r_2^2$ = total surface area of grains per unit volume.

$V = \frac{4}{3}\pi N r_3^3$ = total volume of grains per unit volume of space.

ρV = total mass of grains per unit volume of space.

$\delta = \rho V(\mu n_0 H)^{-1}$ = fraction of total material in the form of dust.

We define β, the rate-coefficient for the formation of H_2 by

$\beta n_0 n_1$ = number of H_2 molecules formed per unit volume per unit time.

The number of hydrogen atoms impinging upon the grains per unit volume per unit time is

$$n_1 A\left(\frac{H}{2\pi kT}\right)^{3/2}\int_0^\infty\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} u\exp\left[-\frac{H}{2kT}(u^2+v^2+w^2)\right]du\,dv\,dw = n_1 A\left(\frac{kT}{2\pi H}\right)^{1/2} \tag{5.1}$$

where k is Boltzmann's constant. Here we have, of course, neglected the speeds of random motion of the grains compared with those of the atoms.

Using the probability ϖ defined in Section 4, the number of hydrogen molecules formed per unit volume per unit time is then from (5.1)

$$\beta n_0 n_1 = \tfrac{1}{2}\varpi n_1 A\left(\frac{kT}{2\pi H}\right)^{1/2} = \frac{3}{2}\left(\frac{kHT}{2\pi}\right)^{1/2}\frac{\varpi\mu\delta}{2\rho}\frac{r_2^2}{r_3^3}n_0 n_1. \tag{5.2}$$

6. *Properties of grains.*—We regard the grains as effectively spherical in form. Oort and van de Hulst (1946) have given a theory of their growth that leads to a distribution-function $N(r)dr$ for the number having radius r to $r+dr$. The function $N(r)$ is of a complicated character, but we have found from the numerical values given by these authors that a good approximation over the significant range of r is given by

$$N(r) = C\exp\left[-(r/a)^p\right] \tag{6.1}$$

where C, a, p are constants.

Noting that

$$\int_0^\infty r^q \exp\left[-(r/a)^p\right]dr = \frac{a^{q+1}}{p}\Gamma\left(\frac{q+1}{p}\right),$$

we then find on using the appropriate values of q

$$N(r)dr = \frac{Np}{a\Gamma(p^{-1})}\exp\left[-(r/a)^p\right]dr \tag{6.2}$$

$$r_1 = a\Gamma(2p^{-1})/\Gamma(p^{-1})$$

$$r_2^2 = a^2\Gamma(3p^{-1})/\Gamma(p^{-1})$$

$$r_3^3 = a^3\Gamma(4p^{-1})/\Gamma(p^{-1}).$$

Thus

$$\frac{r_2^2}{r_3^3} = \frac{\kappa(p)}{r_1}, \tag{6.3}$$

where

$$\kappa(p) = \frac{\Gamma(2p^{-1})\Gamma(3p^{-1})}{\Gamma(p^{-1})\Gamma(4p^{-1})}. \tag{6.4}$$

Substituting in (5.2) we then have

$$\beta = \tfrac{3}{2}\left(\frac{kHT}{2\pi}\right)^{1/2}\frac{\varpi\mu\delta}{\rho r_1}\kappa(p). \tag{6.5}$$

Inserting numerical values of the general physical constants, this may conveniently be written in the form

$$\beta \doteqdot 9{\cdot}1\times 10^{-15}\frac{\varpi\mu\delta}{(10^5 r_1)\rho}\left(\frac{T}{100}\right)^{1/2}\kappa(p). \tag{6.6}$$

7. *Numerical values.*—As a basis for discussion, we are going to specify what we shall take as "favourable" conditions for the operation of the process. We consider each factor of β in (6.6), stating the value we shall use followed by the reasons for doing so.

$\varpi = 1$. The formation of H_2 by the process considered here has been studied experimentally by Wood and Wise (1958) for the case of metal surfaces. They found values of ϖ in the range 0·10 to 0·25 (except in the case of aluminium for which a much lower value was obtained). These values were for room temperatures. The formation of H_2 on non-metallic surfaces was studied by Smith (1943) in the work already quoted. He found values about 0·10

for a number of substances, and, of course, much lower values for substances like glass and quartz that hardly come in question here. But it should also be noticed that the experiments as conducted depend on the fact that about the maximum value $\varpi = 1$ is realized for the catalytic surface on the probe, this being usually platinum. Thus the situation is that in the laboratory the maximum value is attainable while values of over one-tenth of the maximum are not uncommon.

Now a grain in interstellar space would have a surface that is chemically much cleaner than any that could be prepared in the laboratory. Further, a grain that has grown in interstellar gas would be expected to be more porous than the surfaces used in the laboratory experiments just mentioned. This feature would increase the chance that an impinging atom will encounter an active site in the surface, even though it does not affect the calculation of the number of impinging atoms. The latter depends upon the area of a circumscribed sphere and not upon the area when the rugosity of the surface is taken into account.

At the moment we are asking what the process *could* do in favourable conditions. For the purpose of discussion we shall use the maximum value $\varpi = 1$ on the ground that there are quite a number of real substances out of which the grains might be made that could give nearly this value. Indeed Kahn (1955) quotes a conclusion of J. H. de Boer that most probably $\varpi = 1$ and treats it as so significant that it may be necessary to reconcile all astrophysical results with this value.

$2{\cdot}5 < p < 3{\cdot}0$; $\kappa(p) = 0{\cdot}55$. If we seek to represent the calculations of Oort and van de Hulst (1946) as summarized in their Table 2 by a law of the form (6.1) we find $2{\cdot}5 < p < 3{\cdot}0$. The values of the Γ-functions in (6.4) give $\kappa(2{\cdot}5) \doteqdot 0{\cdot}54$, $\kappa(3{\cdot}0) \doteqdot 0{\cdot}56$. Thus it is a sufficiently good approximation to take $\kappa = 0{\cdot}55$ without specifying the value of p more precisely. Van de Hulst (1948) gives the estimate $p = 2{\cdot}6$.

The fact that the particular factor κ is not very sensitive to the distribution of the grain-radii does not mean that the whole result is insensitive to this. For the calculation of r_1 also depends upon the distribution law. In fact, the work of Oort and van de Hulst shows that their distribution-function yields a value of r_1 for all the grains which is a good deal smaller than the mean radius of those grains mainly responsible for observed optical effects.

$r_1 = 10^{-5}$ cm. The authors referred to obtained $r_1 \doteqdot 1{\cdot}5 \times 10^{-5}$ cm. But the mean radius for metallic grains is usually quoted as $r_1 \doteqdot 10^{-5}$ cm; since we are considering what could happen under favourable conditions, it is more natural to use this value. Moreover, the actual grains are certainly not spherical, since spherical grains could not produce the observed polarization of light passing through the interstellar medium. A non-spherical grain has a greater area for collisions than a spherical one of the same volume and this is equivalent to reducing the effective value of r_1 in (6.5).

$\rho = 1{\cdot}1$ g cm^{-3}. This is the value quoted by Allen (1955). It would be small for a compact metallic grain but, as we have said, a grain is more likely to be a porous structure. So this value may still be appropriate even if the density of the material of the grain is somewhat greater.

$\delta = 10^{-2}$. This seems to be about the value most commonly inferred. The mean figures quoted by Allen (1955) would give a value 1·4 times greater. The value cannot be much more than this because there is almost certainly not a large

enough proportion of heavy elements to provide the material that would be required. On the other hand, some estimates are appreciably less, but here we must consider ourselves to be dealing with regions where grains are relatively plentiful.

$\mu = 1{\cdot}5$. The total number of hydrogen atoms per unit volume n_0 is also approximately the total number of all atoms per unit volume. So μ as defined is effectively the mean atomic weight of the particles; the heavier atoms contribute appreciably to the total mass but not much to the total number. With this interpretation, μ is usually estimated to be about 1·5 (Ebert 1955, McCrea 1957).

$T = 100$ degrees. This is a typical value for a H I region. A small change in T would not make much difference since T occurs only to the power $\frac{1}{2}$.

Inserting all these values in (6.5) we get

$$\beta = 6{\cdot}8 \times 10^{-17} \text{cm}^3\text{s}^{-1}. \tag{7.1}$$

The conditions described in this section will be called "favourable" for present purposes. The chief way in which we have given the process the best possible chance is by using the maximum value for ϖ. But, as we have pointed out, this amounts only to assuming that the projected area of a grain has the maximum possible activity; if the surface is very porous, this need not imply that the actual surface achieves maximum activity. Apart from this, we have only taken conditions that appear from reported studies to be those actually found in at any rate some H I regions.

The result (7.1) cannot be described as an upper bound for β, that is to say, the conditions cannot be said to be the *most* favourable that could occur. In particular, if there happened to be very many more grains of very small radius, or if the grains were of a thread-like texture, the active surface for a given mass of dust could be considerably greater than we have taken it to be.

Obviously we cannot give a meaningful overall lower bound for β. But we can say that if all the conditions are about those we have assumed, except only in regard to the chemical activity on the surface of the grains, then β is unlikely to be reduced below about 10^{-4} times the value in (7.1) for any possible composition of the grains. For the experimental results on the formation of H_2 do give values for some substances as much as 10^{-4} times less than the maximum, but none smaller than this except for surfaces like those of glass and quartz which presumably can be ruled out of the present considerations.

8. *Formation of CH.*—If the molecule CH is formed by a similar process the rate of formation per unit volume per unit time would be expressible as

$$\beta_{\text{CH}} n_0 n_{\text{C}} \tag{8.1}$$

where n_{C} is the number of carbon atoms per unit volume. We should have

$$\frac{\beta_{\text{CH}}}{\beta} = 2\left(\frac{H}{C}\right)^{1/2} \frac{\varpi_{\text{C}}}{\varpi}. \tag{8.2}$$

Here C is the mass of the carbon atom and the corresponding factor in (8.2) takes account of the fact that the mean speed of the atoms is inversely proportional to the square root of the mass. The quantity ϖ_{C} is the probability that a carbon atom striking the grain will produce a CH molecule.

The greatest possible value of ϖ_C is unity. For, since hydrogen is so much in excess of carbon the surface layer would consist of hydrogen rather than carbon and so, in a favourable case, a carbon atom impinging on a grain would find an almost complete layer of hydrogen present. The case of H_2 was different because half the impinging atoms were needed to replenish the layer; this is why the factor 2 appears in (8.2).

With these values, we find for "favourable" conditions

$$\beta_{CH} = 3{\cdot}9 \times 10^{-17} \text{cm}^3\text{s}^{-1}. \tag{8.3}$$

Here we have treated the grains as being electrically uncharged. A charge could make a significant difference in the case of CH, but probably not in the case of H_2. This possibility will be discussed elsewhere.

9. *Collisions between clouds.*—The foregoing results concern a region of undisturbed interstellar matter. If two clouds of the material collide, the immediate effect in general will be that a shock wave will move into each cloud from the interface, since the relative velocities of clouds are in general supersonic. In the region between the two shock-fronts the material will be at a higher density and kinetic temperature than in the undisturbed regions.

We may recall the simplest shock conditions (as stated, for example, by McCrea (1956) for a somewhat similar problem). Let c_i, ρ_i, T_i, u_i be the sound–speed, density, kinetic temperature and Mach-number of the flow, where $i = 1$ on the supersonic side of the shock and $i = 2$ on the subsonic side; let γ be the ratio of the specific heats. Then

$$2\gamma u_1^2 u_2^2 - (\gamma - 1)(u_1^2 + u_2^2) - 2 = 0, \tag{9.1}$$

$$[2 + (\gamma - 1)u_2^2]c_2^2 = [2 + (\gamma - 1)u_1^2]c_1^2, \tag{9.2}$$

$$\rho_2 c_2 u_2 = \rho_1 c_1 u_1. \tag{9.3}$$

If u_1 is considerably greater than unity, from (9·1) we have approximately

$$u_2^2 = \frac{\gamma - 1}{2\gamma}. \tag{9.4}$$

Then (9.2) gives

$$\frac{T_2}{T_1} = \left(\frac{c_2}{c_1}\right)^2 = \frac{2\gamma(\gamma - 1)}{(\gamma + 1)^2} u_1^2, \tag{9.5}$$

and (9.3) gives

$$\frac{\rho_2}{\rho_1} = \frac{\gamma + 1}{\gamma - 1}. \tag{9.6}$$

Equation (9·6) gives the maximum compression that can be produced behind a shock. Equation (9.5) could be got from the conversion of energy of flow into thermal energy that results from the shock.

If similar clouds collide with relative speed $2v$ then, from the definitions of u_1, u_2, we have

$$v = c_1 u_1 - c_2 u_2 \tag{9.7}$$

which gives, using (9·3), (9.6)

$$c_1 u_1 = \tfrac{1}{2}(\gamma + 1)v. \tag{9.8}$$

If one of the clouds has thickness l in the direction of motion, then the time for the shock wave to traverse the cloud is

$$t_l = \frac{l}{c_1 u_1} = \frac{2l}{(\gamma+1)v}. \tag{9.9}$$

We have also

$$c_i^2 = \gamma \mathfrak{R} T_i/\mu \tag{9.10}$$

where μ is the mean atomic weight and $\mathfrak{R}$ is the gas-constant.

In deriving these results we have assumed that the gas does not lose energy. On the other hand, if the shocked material does cool rapidly by radiation to temperature $T_2'(< T_2)$ then, in the present case where u_1 is taken to be considerably greater than unity, it can be shown that the density ρ_2' is given approximately by the simple relation

$$\rho_2' T_2' = \rho_2 T_2, \tag{9.6'}$$

where ρ_2, T_2 are given by (9.5), (9.6).

In our case we take $\gamma = 5/3$ as for a monatomic gas and we take $\mu = 1{\cdot}5$ as before. Then (9·6), (9·5), (9·10) become

$$\rho_2 = 4\rho_1 \tag{9.11}$$

$$(T_2/T_1)^{1/2} = c_2/c_1 \doteqdot 0{\cdot}56 u_1 \tag{9.12}$$

$$c_1 \doteqdot 0{\cdot}96 (T_1/100)^{1/2} \text{ km s}^{-1}. \tag{9.13}$$

Then if we take as what is probably a typical set of values

$$T_1 = 100 \text{ degrees}, \quad l = 10 \text{ parsecs}, \quad v = 7 \text{ km s}^{-1}, \tag{9.14}$$

we get approximately

$$c_1 = 1 \text{ km s}^{-1}, \quad u_1 = 28/3, \quad (T_2/T_1)^{1/2} = c_2/c_1 = 5{\cdot}2, \tag{9.15}$$

$$t_l \doteqdot 1{\cdot}1 \text{ million years}. \tag{9.16}$$

Since the rates of the reactions in which we are interested are proportional to the square of the density and the square root of the temperature, we see that (in this example) these rates would be enhanced by a factor $4^2 \times 5{\cdot}2 \doteqdot 83$ as a result of the collision and that the time during which such enhancement would be effective would be at any rate of the order of 10^6 years. Moreover, if the reactions do not proceed at the maximum rate, it is known from experiment that the probability corresponding to ϖ increases with increasing temperature in the ranges studied. All this is under the assumption that the collision between the clouds is not so violent that the dust grains are destroyed.

If there is appreciable cooling by radiation so that the shocked material is characterized by ρ_2', T_2' instead of ρ_2, T_2 then the reaction-rate is to be further multiplied by a factor

$$(\rho_2'/\rho_2)^2 (T_2'/T_2)^{1/2}.$$

If (9.6′) is applicable, we see that this is approximately

$$(T_2/T_2')^{3/2}$$

which is always greater than unity. Thus in estimating the enhancement in the reaction-rate produced by a collision we get the minimum rate if we neglect radiation. We owe this observation to Dr F. D. Kahn.

10. *Characteristic times.*—If the rate of dissociation of molecules, once they are formed, is negligible, then the definition of β gives

$$dn_1/dt = -2\beta n_0 n_1. \tag{10.1}$$

The factor 2 enters because two free atoms disappear when a molecule is formed. From (10.1) the time to reduce n_1 from n_0 to any smaller value is

$$t = t_1 \ln (n_0/n_1), \tag{10.2}$$

where

$$t_1 = (2\beta n_0)^{-1}. \tag{10.3}$$

The time t_1 is that required to convert a fraction $(1 - e^{-1})$ of the atoms into molecules and it may be taken as the *characteristic time* for H_2 formation. It should be noted from (10·2) that the time required to convert a small fraction ϵ of the atoms into molecules is approximately ϵt_1.

The only dissociation process in interstellar space that need be considered is photo-dissociation by starlight. Let α be the chance that an H_2 molecule is dissociated in unit time. Then if molecule formation is negligible we have

$$dn_2/dt = -\alpha n_2. \tag{10.4}$$

Thus the characteristic time for H_2 dissociation is

$$t_2 = \alpha^{-1}. \tag{10.5}$$

If molecule formation and dissociation are both occurring and if there is a steady state, we have

$$\beta n_0 n_1 = \alpha n_2 \quad (n_0 = n_1 + 2n_2) \tag{10.6}$$

or

$$\frac{n_1}{2t_1} = \frac{n_2}{t_2}. \tag{10.7}$$

If there is not a steady state we have

$$dn_1/dt = -(2\beta n_0 + \alpha) n_1 + \alpha n_0. \tag{10.8}$$

This shows that the characteristic time to approach a steady state is $(2\beta n_0 + \alpha)^{-1}$. This is never very different from the *shorter* of the times t_1, t_2.

In the case of CH there are minor modifications because only one carbon atom goes into a molecule. If $\beta_C, \alpha_C, t_C, t_{CH}$ correspond to β, α, t_1, t_2 we find

$$t_C = (\beta_C n_0)^{-1} \tag{10.9}$$

$$t_{CH} = \alpha_C{}^{-1} \tag{10.10}$$

and for a steady state

$$\beta_C n_0 n_C = \alpha_C n_{CH} \quad \text{or} \quad \frac{n_C}{t_C} = \frac{n_{CH}}{t_{CH}}. \tag{10.11}$$

11. *Discussion*: *CH.*—We discuss the case of CH first because more is known about it than about H_2.

Bates (1951) calculated a rate-coefficient for the formation of CH by radiative association which for the present purpose may be quoted as 4×10^{-18}, which is the mean of his two values depending upon the state of the carbon atoms. Bates and Spitzer (1951) concluded that this is too small by a factor of 500 to 1000 to account for the observed amount of interstellar CH. The value of β_{CH} in (8·3) is about 10 times the quoted value for radiative association. So we

have to conclude that even in favourable conditions the process considered here will not account for the required amount of CH in the interstellar material considered by Bates and Spitzer.

However, if this same material takes part in a cloud-collision of the sort considered in Section 9, then the rate of formation of CH by the process contemplated would be enhanced by another factor about 80 during the time that the effect of the collision persists. Thus in all we can obtain a rate of about 800 times the rate of radiative association in the undisturbed gas.

It is seen, therefore, that the process considered would account for about the observed amount of CH in some regions provided the conditions are about what we have called "favourable" to the process. Since these are regions of enhanced temperature, they offer a good chance that the conditions are in fact favourable.

Since we have estimated that the conditions could persist for about 10^6 years, and since it is usually estimated that an interstellar cloud takes part in a collision once in about 10^7 years, the regions concerned could include something of the order of 10 per cent of all the cloud material.

It is to be noticed that the rate of radiative dissociation of CH calculated by Bates and Spitzer makes t_{CH} a few thousand years. So molecules produced during the "enhanced" conditions would decay quite rapidly after such conditions ceased to obtain. Moreover, this figure makes it very evident that the problem of CH is certainly that of its production *in situ* in interstellar material; there can be no question of the observed CH having survived after having been injected into the material from elsewhere.

Bates and Spitzer have examined the problem with great care and have stressed the extreme difficulty of accounting for the existence of CH. Consequently the present results do appear to be significant. They demand only processes that we know must occur i.e. collisions between clouds and clouds and collisions between atoms and grains, together with a known type of surface reaction. The only uncertainty, admittedly a crucial one, is whether the particular reaction concerned is fairly efficient or quite ineffectual.

It is important to remark that if the presence of CH is to be accounted for at all by the process suggested it probably *has* to be associated with collisions between clouds. For, if the process could account for the CH apart from such collisions, then, since collisions certainly do occur and since the process would certainly proceed at an enhanced rate as a result, we should have an embarrassingly high rate of production in colliding clouds.

It seems therefore to be a justifiable provisional inference that CH is to be accounted for in the manner suggested, and, in particular, that the surface reaction has the required efficiency. No alternative suggestion has been offered (except the possible evaporation of CH_4 from the grains mentioned by Bates and Spitzer which may produce CH in regions other than those with which we are concerned). Bates and Spitzer did discuss the possibility of reactions on grains, but without special reference to the effects of cloud-collisions upon rates of reaction.

Actually we are not here concerned primarily to account for the observations of CH but rather to test the hypothesis concerning the significance of surface reactions in order to be in a better position to discuss the case of H_2.

We can in fact say of the suggested process (*a*) that it appears adequate to account for the observed amount of CH; (*b*) that, if it does, it operates with

fairly high efficiency and, conversely, the supposition that it operates with high efficiency would not lead to absurd conclusions; (*c*) that if it operates in the case of CH, it almost certainly operates in the case of H_2.

It should be added that Bates and Spitzer found that the use of an empirical oscillator strength instead of the computed one actually employed would have reduced the discrepancy of 500 to 1000 mentioned above by a factor of about 30, but this has no effect on our conclusions.

12. *Discussion*: H_2.—As a basis for discussion, we shall give some figures for what we have called favourable conditions.

Using a recent study of the relevant radiation-field, Kahn (1955) estimated for the quantity α in Section 10 the value

$$\alpha \doteqdot 3{\cdot}3 \times 10^{-14}\,\mathrm{s}^{-1} \quad (t_2 \doteqdot 10^6 \text{ years}). \tag{12.1}$$

(As he says, an earlier study of the radiation-field would make $\alpha \doteqdot 1{\cdot}1 \times 10^{-14}\,\mathrm{s}^{-1}$.)

The condition for a steady state (10·6) may be written

$$\frac{n_1}{n_0} = \frac{\alpha}{\alpha + 2\beta n_0}, \qquad \frac{n_2}{n_0} = \frac{\beta n_0}{\alpha + 2\beta n_0} \tag{12.2}$$

giving, respectively, the fractions of the hydrogen in the form of free atoms and of molecules. Using the value of α in (12.1) and the value of β in (7.1) adjusted for the various temperatures T, we obtain the following table to show some typical values.

TABLE I

Steady-state composition of interstellar hydrogen: n_1 free atoms, n_2 molecules cm^{-3}.

n_0 / T	1		10		10^2		10^3		∞	
deg.	n_1	n_2	n_1	n_2	n_1	n_2	n_1	n_2	n_1 ($\alpha/2\beta$)	$n_0 t_1$ years
50	1·0	0·001	9·7	0·14	77	11	255	372	343	$3{\cdot}3 \times 10^6$
100	1·0	0·002	9·6	0·20	71	15	195	402	243	$2{\cdot}3 \times 10^5$
500	0·99	0·005	9·2	0·42	52	24	98	451	109	$1{\cdot}0 \times 10^5$

The values of $n_0 t_1$ from (10.3) are also shown. But it should be remembered from Section 10 that the significant time is approximately the smaller of t_1, t_2. So we see that the states shown in Table I would all be attained in times between about 10^5 and 10^6 years, and these are quite short times for most processes in interstellar material. That is to say, if the material has the assumed properties and is exposed to the assumed radiation field, it will in general be observed to show the characteristics of a steady state.

It is unfortunate that molecular hydrogen cannot be observed directly by any means at present available. In the circumstances, probably the best test of the figures is along the lines of Kahn's (1955) discussion of the cooling effect of H_2. He points out that, for cooling the material of interstellar clouds after collisions, H_2 is very efficient down to about 500 degrees but less efficient below that temperature. Nevertheless he estimates that only 0·05 molecule cm^{-3} would suffice to produce the harmonic mean temperature of 100 degrees which is about what is required by observation.

Actually Kahn uses $n_0 = 20$ as a typical value for a cloud. Referring to the table, it might then appear that the values of n_2 are rather large, compared with the value 0·05 just mentioned. However, further reference to Kahn's discussion

shows that the values of the order of 0·2 which we find for n_2, when n_0 is of the order 10, would not produce a mean temperature much below 100 degrees. For Kahn estimates that the much larger value $n_2 = 23$ would be needed to produce a mean temperature of 50 degrees. Also, it is seen that the value of n_2 is rather sensitive to the value of n_0 at the values of n_0 concerned. Thus, at $T = 100$ degrees, although we have $n_2 \doteqdot 0{\cdot}2$ for $n_0 = 10$, we should have $n_2 = 0{\cdot}05$ for $n_0 \doteqdot 5$. We may conclude that the values of n_2 in the table would not be ruled out by the conditions for cooling although somewhat smaller values would be acceptable. Of course, it is satisfactory that "favourable" conditions do give values of n_2 exceeding the minimum required for cooling. It is to be noted that Kahn's estimate of the value required for cooling does not involve the parameter α.

Another possible test is provided by the very simple consequence of (12.2) that

$$n_1 < \alpha/2\beta \quad (\text{all } n_0). \tag{12.3}$$

The values of $\alpha/2\beta$ are shown in the table. Physically the property expressed by (12.3) is simply that by compressing the gas at any given temperature *the number-density of free atoms can never be made to exceed a certain limit because all the rest of the hydrogen will go into molecular form.*

Now it has always seemed remarkable that observations have never revealed or demanded densities of hydrogen more than about 100 atoms cm^{-3}. It is very tempting to accept the simple explanation offered by the present work. If indeed 100 atoms cm^{-3} is the upper bound in actuality, this would indicate a value of β somewhat above the "favourable" value with the estimate of α used here, or somewhat below the favourable value with the older estimate of α. This would then be in satisfactory accord with the requirements for cooling. It is to be noted that this possible test does depend on α.

Here we are saying that the density of interstellar matter may become great in certain regions but that this will not be revealed by the density of atomic hydrogen. This explanation of the absence of observations of high densities of atomic hydrogen might be criticized because it might be expected that regions of high overall density, if they exist, would have been detected by the behaviour of *some* constituent, even if not by atomic hydrogen. Actually, further discussion is needed in order to say whether an increase in density of any other particular constituent would or would not be concealed as in the case of hydrogen. Besides, a region of very high density is likely to be of relatively small extent, simply because it has to be produced by the compression of normal interstellar material, and so it may merely escape detection for this reason. Hence at present not too much stress should be placed upon this second possible test by itself. Taken in conjunction with the first, however, it may have some significance.

13. *Conclusion.*—The following features of interstellar matter could apparently all be accounted for if the conditions for molecule formation are about what we have called "favourable":

(*a*) The observed quantity of CH.

(*b*) The cooling of interstellar clouds after collisions.

(*c*) The lack of observed densities of atomic hydrogen exceeding about 100 atoms cm^{-3}.

If these results are significant, whatever be the difference between actual conditions and favourable conditions could easily be accounted for by uncertainties in the values of any or all of the parameters involved. In particular, there is no reason to suppose that most of the uncertainty is in the probability ϖ; the inference would be, indeed, that ϖ cannot be much less than unity.

The importance of this, if correct, is that if cold interstellar gas undergoes compression as a start to the process of star-formation, then the gas will be mostly in molecular form at any rate after a density equivalent to about 1000 atoms cm^{-3} has been attained, i.e. after cloud material has been compressed by a factor of about 100. It must be remembered that the compression required in order to produce stars is by a factor of about 10^{23}. So we see that the problem of star formation is effectively that of forming stars when starting with cold *molecular* hydrogen.

After all this discussion, however, it seems that this last conclusion must be valid whatever the process of molecule-formation! For the case of CH shows that radiative association is not likely to fall behind our "favourable" process by a factor of more than about 10^3. Also, as stated at the end of Section 7, experiment shows that the most inefficient surface reactions are not likely to be worse than the favourable case by a factor of more than about 10^4. Thus at the worst, the effective value of the parameter β could scarcely be reduced by more than 10^4, and even then we should expect to have the hydrogen mostly in molecular form when the cloud material has been compressed by a factor not exceeding 10^6. The difference between 10^2 and 10^6 is not of great importance when compression by 10^{23} is needed. This conclusion is likely to be important in theories of star formation (McCrea 1960).

Royal Holloway College,
Englefield Green,
Surrey:
1960 *March* 21.

References

Allen, C. W., 1955, *Astrophysical Quantities* (London).
Bates, D. R., and Spitzer, L., 1951, *Ap. J.*, **113**, 441.
Bates, D. R., 1951, *M.N.*, **111**, 303.
Ebert, R., 1955, *Zs. f. Astrophys.*, **37**, 217.
Herzberg, G., 1955, *Mem. Soc. R. Sc. Liège* (4), **15**, 291.
van de Hulst, H. C., 1948, *Harvard Monograph*, No. 7, 73.
Kahn, F. D., 1955, *Mem. Soc. R. Sc. Liège* (4), **15**, 578.
McCrea, W. H., 1956, *Ap. J.*, **124**, 461.
McCrea, W. H., 1957, *M.N.*, **117**, 562.
McCrea, W. H., 1960, *Proc. Roy. Soc.*, A, **256**, 245.
Oort, J. H., and van de Hulst, H. C., 1946, *B.A.N.*, **10**, 187.
Smith, W. V., 1943, *J. Chem. Phys.*, **11**, 110.
Wood, B. J., and Wise, H., 1958, *J. Chem. Phys.*, **29**, 1416.

PHOTOHELIOGRAPHIC RESULTS 1956*

[*Greenwich Photoheliographic Results* have hitherto formed part of the annual volume entitled *Greenwich Observations*. The last issue of this volume, and of its separate parts, refers to the year 1955. As from 1956, the solar results obtained at the Royal Greenwich Observatory will be published annually, in somewhat shortened form, in a new series of *Bulletins** issued by H.M. Stationery Office.

This annual *Bulletin* will also supersede the series of short papers hitherto appearing in *Monthly Notices* under the title " Mean daily areas and heliographic latitudes of sunspots ". The last of this series, for the year 1955, appeared in *M.N.*, **116**, 486, 1956.]

Summary

Daily solar photographs taken in integrated light at Herstmonceux (supplemented where necessary by plates from the Cape or from Kodaikanal) are measured to yield the following results, presented in tabular form:—

(1) Positions and areas of sunspots for each day of the year (pp. C8–C61). The positions are referred alternatively to a system of apparent polar coordinates on the Sun's disk and to a system of heliographic coordinates. The areas are given for umbrae and whole spots and are corrected for foreshortening.

(2) A general catalogue of groups of sunspots, whether lasting two days or more (pp. C64–C75) or seen on one day only (pp. C76–C78). This catalogue enables the history of a given group to be traced throughout its appearance on the disk.

(3) Summary tables (pp. C79–C85) of total areas of sunspots and faculae for each day of the year, of mean areas of sunspots and faculae and of mean heliographic latitudes of sunspots for each rotation and for the year.

A short review of the chief features of the solar record for 1956 is included.

* The full text of this paper is published as *Royal Greenwich Observatory Bulletin* No. 14, 1959.